Pós-Graduação em Ciência da Computação

"Evaluation of the Use of Augmented Reality Tools in the Education Field"

por

Manoela Milena Oliveira da Silva

Dissertação de Mestrado

Universidade Federal de Pernambuco
posgraduacao@cin.ufpe.br
www.cin.ufpe.br/~posgraduacao

RECIFE
2015

CENTRO DE INFORMÁTICA

PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

Evaluation of the Use of Augmented Reality Tools in the Education Field

MANOELA MILENA OLIVEIRA DA SILVA

Este trabalho foi apresentado à Pós-Graduação em Ciência da Computação do Centro de Informática da Universidade Federal de Pernambuco como requisito parcial para obtenção do grau de Mestre em Ciência da Computação

Orientado pela Profa. **Veronica Teichrieb**

Co-orientado pela Profa. **Patricia Smith Cavalcante**

RECIFE
2015

Catalogação na fonte
Bibliotecária Jane Souto Maior, CRB4-571

S586e Silva, Manoela Milena Oliveira da
Evaluation of the use of augmented reality tools in the education field / Manoela Milena Oliveira da Silva. – Recife: O Autor, 2015.
140 f.: il., fig.

Orientador: Veronica Teichrieb.
Dissertação (Mestrado) – Universidade Federal de Pernambuco. CIn, Ciência da Computação, 2015.
Inclui referências e apêndices.

1. Realidade aumentada. 2. Educação. I. Teichrieb, Veronica (orientadora). II. Título.

006.8 CDD (23. ed.) UFPE- MEI 2015-174

Dissertação de Mestrado apresentada por **Manoela Milena Oliveira da Silva** à Pós Graduação em Ciência da Computação do Centro de Informática da Universidade Federal de Pernambuco, sob o título "**Evaluation of the Use of Augmented Reality tools in the Education Field**" orientada pela **Profa. Veronica Teichrieb** e aprovada pela Banca Examinadora formada pelos professores:

_____

Profa. Patricia Cabral de Azevedo Restelli Tedesco
Centro de Informática/UFPE

_____

Profa. Veronica Gitirana Gomes Ferreira
Centro de Educação /UFPE

_____

Profa. Veronica Teichrieb
Centro de Informática / UFPE

Visto e permitida a impressão.
Recife, 31 de agosto de 2015.

_____

Profa. Edna Natividade da Silva Barros
Coordenadora da Pós-Graduação em Ciência da Computação do
Centro de Informática da Universidade Federal de Pernambuco.

ACKNOWLEDGEMENTS

I would like to thank my advisors Veronica Teichrieb and Patricia Smith for their care, patience and support throughout this entire process. Not only they provided all the guidance I needed, but they inspired me to be a better professional.

I also need to thank the Voxar Labs family. Being part of a group with these passionate and talented professionals was very important to me. Their support was very important and it can be seen throughout this work.

I am also thankful for the friends I did at Edumatec. I was very glad to have attended those disciplines and have met such great people that helped me a lot.

I am also indebted to the schools, teachers and students who kindly and generously participated in this study.

Above all, I would like to express my deepest gratitude to God, my Family and my dear friends who gave me all the support I needed. They encouraged me to be a better person and professional. Moreover, they inspire me to always give my best. I am also grateful for my beloved husband, Rafael, who has always been supportive and inspiring. His contributions are spread throughout this work.

ABSTRACT

Augmented Reality technology (AR) has a huge potential to be applied in the education field. The coexistence of real and virtual environments enables experiences that would not be possible without this technology. Some of the reasons why AR learning experiences differ from other technology are: (i) it enables contextualized interaction between real and virtual worlds, (ii) it enables tangible interaction metaphors for object manipulation, and, finally, (iii) it enables smooth transition between the real and virtual contents. While AR offers new learning opportunities, it also creates new challenges for education in different domains, such as technological, learning and pedagogical issues. This work intends to provide some reflections about the challenges involved in the process of evaluating AR educational technologies. In order to better understand those issues, a systematic review was carried out aiming to identify how AR technology has been evaluated. Taking into account lessons learned during the review, a projective educational AR tool, especially designed to young children education, the ARBlocks, was evaluated. This tool was evaluated in the field of language learning with three different groups. The study involved the teacher as an instructional designer along with the use of multiple metrics. From the analysis of the ARBlocks in the classrooms, it was possible to observe that this tool offered different possibilities for language teaching to young children. The results obtained demonstrated that, in general, the ARblocks contributed to student's learning and the practice and reinforcement of language abilities. From the reflections presented, some guidelines were proposed in order to assist the evaluation of AR educational tools. The use of multiple metrics as well as the active involvement of teachers in the elaboration of contents are encouraged as way to better understand the impact of technology in the teaching and learning process.

Keywords: Augmented Reality. Education. Evaluation. Educational Systems.

RESUMO

A tecnologia de Realidade Aumentada (RA) possui grande potencial de aplicação na área educativa. A coexistência de ambientes reais e virtuais abre possibilidades de aprendizado que não poderiam ser possíveis sem este tipo de tecnologia. Algumas razões pelas quais experiências de aprendizagem com RA se diferem das demais são: (i) suporte a uma interação contextualizada entre ambientes reais e virtuais, (ii) uso de metáforas com interfaces tangíveis para manipulação de objetos e, por fim, (iii) a habilidade de transição suave entre o real e o virtual. Enquanto oferece inúmeras novas oportunidades de aprendizagem, a introdução de novas tecnologias com RA cria desafios em diferentes domínios, como o tecnológico, de aprendizagem e os desafios pedagógicos. Este trabalho visa prover algumas reflexões acerca dos desafios envolvidos no processo de avaliação de tecnologias educativas com RA. Como forma de melhor compreender essas questões, foi realizada uma revisão sistemática com vistas a identificar como são realizadas avaliações de ferramentas educacionais com RA. A partir da revisão empreendida, foram observadas as principais formas de avaliação de ferramentas educativas envolvendo esta tecnologia. A partir destes conhecimentos adquiridos, foi realizada a avaliação de uma ferramenta educativa baseada em RA projetiva, especialmente desenvolvida para o ensino infantil, o ARBlocks. Tal ferramenta foi avaliada no campo da aprendizagem de novas línguas com três grupos distintos. O estudo envolveu o professor como *instructional designer*, bem como o uso de múltiplas métricas de avaliação. A partir da análise do ARBlocks em sala de aula, foi possível observar que esta ferramenta oferece inúmeras possibilidades para o ensino de línguas para crianças pequenas. Os resultados obtidos demonstraram que, no geral, o ARBlocks auxiliou no processo de aprendizagem dos estudantes e na prática e reforço das habilidades linguísticas. A partir das reflexões apresentadas, alguns *guidelines* foram propostos com vistas a auxiliar na avaliação de ferramentas educativas com RA. O uso de múltiplas métricas e o envolvimento ativo dos professores na elaboração dos conteúdos é encorajado como forma de melhor compreender os impactos provocados pela tecnologia no processo de ensino-aprendizagem.

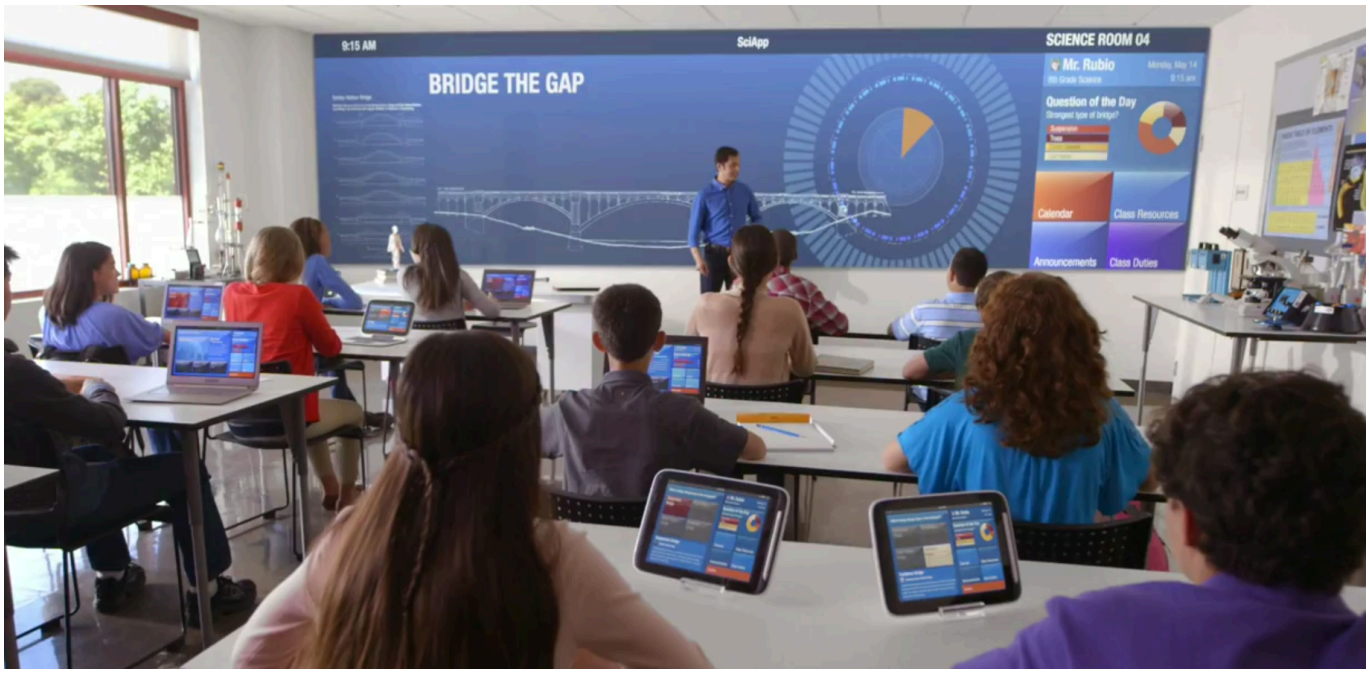
Palavras-chave: Realidade Aumentada. Educação. Avaliação. Sistemas Educacionais.

TABLE OF FIGURES

TABLE OF CONTENTS

1. INTRODUCTION

Figure 1: Classroom with different technology tools



Technology is widely spread in different areas, such as entertainment, education and many others. It is part of everyday life and, therefore, most of the students are already familiar with it as well as it has been increasingly introduced in the classrooms as illustrated in Figure 1. (KENSKI, 2007) explains that its arrival in schools implies a range of challenges to teachers, students and the pedagogical team since the technology brings a double challenge: to adapt schools to its advances and to guide the people involved to master critically this new media.

One piece of technology that has a huge potential to be a valuable educational tool is augmented reality or simply AR (CONSORTIUM, T. N. M., 2011). This technology consists of adding virtual elements to a real scene (AZUMA, BAILLOT, *et al.*, 2001). AR applications combine virtual and real elements in a coherent way so users cannot differentiate them from the real scene.

The coexistence of virtual and real environments allows learners to experience phenomena that otherwise would be impossible in the real world, interact with two- or three-dimensional synthetic objects in the mixed reality and, therefore, develop important abilities that cannot be developed in other technology learning environments (WU, LEE, *et al.*, 2013).

Many researchers have identified AR's potential to the educational environment. For instance, (SILVA, 2012) listed some educational applications that used it. (RADU, 2012) analyzed 32 works that compare AR tools with non-AR systems. He describes some positive and negative aspects

regarding the use of AR technology in educational applications.

While AR offers new learning opportunities, it also creates new challenges for education in different domains, such as technological, learning and pedagogical issues (KERAWALLA, LUCKIN, *et al.*, 2006).

Developments in AR technology have enabled many technologies to be developed for the educational field. Their impact in the educational setting is of utmost importance. Hence, evaluation of learning outcomes is an important step in order to unveil its potential in education.

1.1. OBJECTIVES

This work intends to provide some reflections about the issues involved in the process of evaluating AR educational technology.

As a way to better understand this issue, a systematic review will be conducted in order to investigate how educational AR technology has been evaluated in different learning settings.

Taking into account lessons learned during the review, a projective AR educational tool, the ARBlocks (ROBERTO, FREITAS, *et al.*, 2013) will be evaluated in the field.

The specific objectives of our study are:

a) to evaluate the use of the ARBlocks in the educational field concerning the linguistic concepts and competencies;

b) discuss possibilities of use for AR tools in the classroom environment as well as reflect about the potential and difficulties involved in the introduction of this new piece of technology in the school environment and how teachers can adapt this new tool into their teaching routines;

c) generate guidelines based on our experience in order to assist researchers when conducting similar evaluations.

1.2. DOCUMENT STRUCTURE

The following chapters of this document are organized as follows: chapter 2 presents an overview of AR especially focusing on the educational field. Chapter 3 discusses issues regarding teaching English to young children. Chapter 4 presents a systematic review conducted in order to investigate how AR applications have been evaluated in the educational scenario. Chapter 5 discusses the methodology adopted in this work, the tool evaluated in the study, the ARblocks, as well as a pilot study conducted with the tool. Chapter 6 focuses on the semi-experimental study in which the ARBlocks was evaluated. Chapter 7 presents and discusses the results of our experiment and, finally, chapter 8 presents the conclusions from our work.

2. AR IN EDUCATION

Figure 2: AR Book

AR is a new technology that has gained a lot of attention in the recent years. It refers to a type of technology that combines the real and virtual world in a coherent way. It superimposes virtual information, such as objects, pictures and text in the real word in a seamless way as illustrated in Figure 2. The Incredebooks allow children to interact with virtual information incorporated in the book's storyline (MERCURY INPRESS, 2015).

(AZUMA, 1997) defines AR as combination of virtual and augmented reality and emphasizes its features or characteristics. For a system to be considered AR, it must fulfill three features: combination of real and virtual worlds, real-time interaction, and accurate 3D registration of virtual and real objects. Later on, some new aspects have been added to this concept, such as simulation, on-line effects, and 2-D perspectives elements (SALMI, KAASINEN and KALLUNKI, 2012). Researchers advocate that understanding AR in a broad sense is more productive for educators and designers since this definition implies that varied technologies could be used to implement AR systems, such as desktop computers, handheld devices, head-mounted displays and so on (BROLL, LINDT, *et al.*, 2008) (JOHNSON, LEVINE, *et al.*, 2010) (LIU, 2009).

Therefore, authors argue that AR should be understood as a concept rather than a type of technology in order to be more productive for researchers, educators and designers (WU, LEE, *et al.*, 2013).

AR is not restricted to particular types of display technology, such as head-mounted displays. Besides, although the sense of sight is the most explored in AR applications, (AZUMA, BAILLOT, *et al.*, 2001) explain that this technology can be potentially applied to all senses.

2.1. TAXONOMIES OF AR

(MILGRAM, TAKEMURA, *et al.*, 1994) established a continuum that is helpful to explain the concept of AR. This continuum covers the different possibilities of mixing real and virtual environments. In one of its extremes, we have the real environment and in the other end, the virtual environment. Between those extremes, there are AR and augmented virtuality as can be seen in the Figure 3.

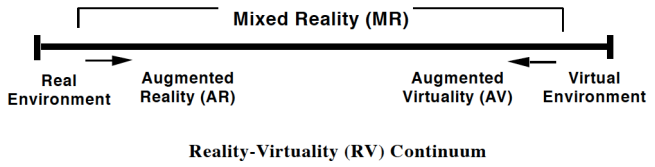


Figure 3: Reality-Virtuality Continuum

In order to provide effective AR experiences, three factors must be taken into account, that are: tracking or registration techniques of the real scenes, display hardware to present the augmented content and, finally, interaction techniques so users can manipulate the augmented contents.

This chapter will detail each of these factors as well as provide some general overview of how AR can be used in a variety of fields, giving special attention to its potential in the field of education.

2.2. TRACKING TECHNIQUES

Tracking or registration is a fundamental enabling technology for AR systems along with displays and graphics. It is also one of the most challenging areas in spatial AR. Not surprisingly, it has been the most popular topic of research of ISMAR conferences from 1998-2008 (ZHOU, DUH and BILLINGHURST, 2008).

Tracking is responsible for "reading" the environment. It is necessary to record the position and orientation of real objects in physical space and to allow consistency between real and virtual objects.

According to (ZHOU, DUH and BILLINGHURST, 2008), tracking techniques can be divided in three different groups that will be appropriately explored below: sensor-based tracking, vision-based tracking and hybrid-tracking techniques.

In Figure 4 is presented an infographic that displays a general taxonomy for AR tracking techniques.

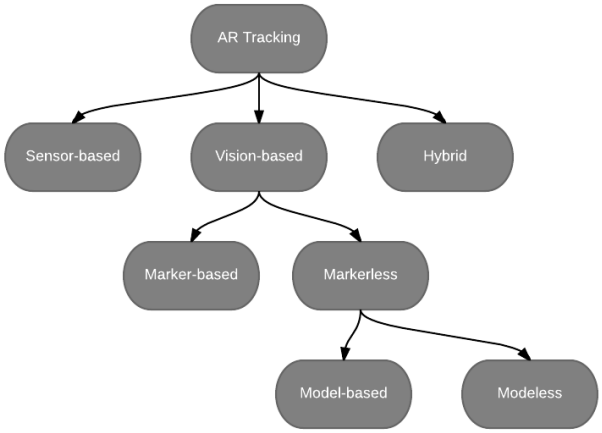


Figure 4: General taxonomy for AR tracking techniques

2.2.1. SENSOR-BASED TRACKING TECHNIQUES

Sensor-based tracking techniques use the devices' sensors to estimate their position and orientation. Current gadgets are embedded with several sensors, such as magnetic, acoustic, inertial, optical and mechanical sensors (ZHOU, DUH and BILLINGHURST, 2008).

Each sensor type has its own advantages and disadvantages which will have more or less impact depending on the system being developed (ROLLAND, DAVIS and BAILLOT, 2001).

2.2.2. Vision-Based Tracking Techniques

Vision-based tracking techniques use image processing and computer vision methods to calculate the camera pose relative to real world objects and, therefore, correctly display graphic content on the scene. It has been the most active area of research in ISMAR conferences (ZHOU, DUH and BILLINGHURST, 2008).

Vision-based tracking techniques can be divided in two groups: marker-based tracking and markerless tracking (RABBI and ULLAH, 2013).

The former approach uses artificial markers to identify objects' position in the scene which reduces computational requirements to calculate camera pose.

There are many marker-based kits available. The most famous is the ARtoolkit (KATO and BILLINGHURST, 1999). Different types of markers include template-based, ID-based, data-matrix, split, frame-markers, QR-codes and random dots (ROBERTO, 2012).

Markerless tracking, on the other hand, uses natural characteristics of the environment such as points, edges, corners and textures to calculate camera pose. In other words, the environment itself acts as a marker.

(RABBI and ULLAH, 2013) point out two different types of markerless tracking: model-based in which a model of an object is used to guide tracking and modeless-based. In the latter, the camera movement is calculated without previous knowledge of the scene.

2.2.3. Hybrid-Tracking Techniques

Many researchers are devoted to explore how different sensors can be combined in order to provide robust tracking with dynamic sensor hand over between tracking technologies. Therefore, different systems may employ hybrid-tracking techniques which combine different sensing technologies to exploit strengths and compensate weaknesses of individual tracking techniques (AZUMA, BAILLOT, *et al.*, 2001).

An example of hybrid-tracking technique is the work of (KURZ and BENHIMANE, 2011) that uses inertial sensors to measure the gravitational force and, therefore, uses this information to improve vision-based tracking for handheld AR applications.

2.3. Display Techniques

(KALKOFEN, SANDOR, *et al.*, 2011) explain that visualization techniques are powerful tools for exploring the real world structures along with additional contextual information.

Along with an accurate tracking system, the display hardware is an important part of an AR system and plays a role when it comes to provide an immersive experience for the user.

Display techniques can be classified according to the user's viewpoint as illustrated in Figure 5.

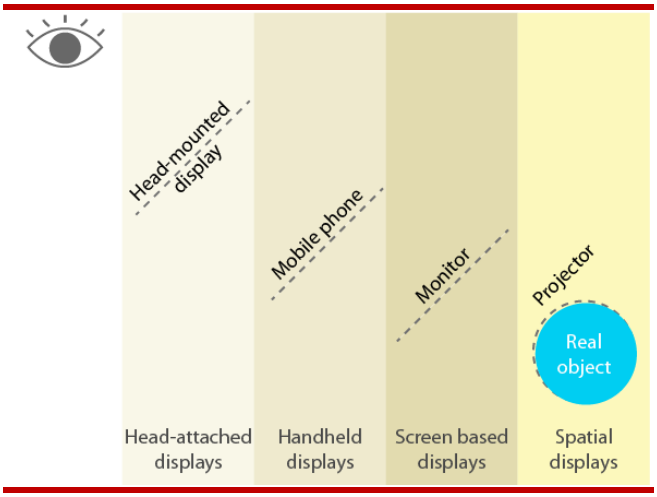


Figure 5: Different forms users can visualize information in AR applications

2.3.1. Head-Attached Displays

Head-attached displays, as the name suggests, are the ones attached to the user's head. Users need to wear them and graphic content is displayed at their sight. Therefore, these displays provide a better field of view (FOV). As for disadvantages, it is important to highlight social and fashion constraints.

Head-attached displays can be divided in three types: retinal displays, head-mounted displays and head-mounted projectors.

Retinal displays exhibit information directly in the user's eyes as presented in (KLEWENO, SEIBEL, *et al.*, 2001). This work uses a virtual retinal display (VRD) to aid low vision users in reading tasks.

Head-mounted projective displays (HMPD) consist of miniature lenses, beamsplitters and displays mounted on the helmet and retro-reflective sheeting materials placed strategically in the environment. In the GO game, users wear a HMPD and manipulate real stone pieces on the virtual board (HUA, GAO, *et al.*, 2001).

Head-mounted displays require users to wear them on their heads via a headband, a helmet or around an eyeglasses frame. They are divided in two different categories: optical see-through (OST) and video see-through (VST).

OST allows the user to see the world with their natural eyes and overlay virtual content onto the users view. An example of this display is the Google Glass (GOOGLE, 2015) as seen in Figure 6.

Figure 6: Google Glass, an example of wearable technology with optical head-mounted display developed by Google company

On the other hand, VST displays obstruct users view of the real world by presenting a video view of the world overlaid by graphics. One example of this type of display is the Oculus Rift (OCULUS VR) as presented in Figure 7.

Figure 7: Oculus Rift is an example of a video see-through display

2.3.2. Handheld Displays

Recent advancements on processing techniques, memory capacity, graphic capabilities and the addition of different sensors, such as build-in camera, accelerometer and magnetometer have

enabled AR applications to operate properly in cellphones and other mobile devices. Those displays are very flexible and allow users to manipulate technology they are already familiar with.

(ZHOU, DUH and BILLINGHURST, 2008) state that handheld displays are a good alternative to AR applications because they are minimally intrusive, socially acceptable, readily available and highly mobile.

As an example of mobile AR applications using handheld displays, we have the CityView AR (HIT LAB NZ) that enables users to visualize how the city of Christchurch-NZ was before an earthquake that happened in September 2010. This application is presented in Figure 8.

Figure 8: By using CityView AR, users are able to point their cellphones to different sites and see how buildings used to be before the earthquake and the demolitions

As for limitations, the small size of the display on handheld devices is not ideal for 3D user interfaces. In addition, it provides less immersion when compared to head-mounted displays. FOV and tracking are also some of its limitations.

2.3.3. Screen Based Displays

Screen based displays use video mixing techniques to display merged images in a regular monitor. (BIMBER and RASKAR, 2006) explain that this technique is usually referred to as “window on the world”.

The mirracle (TUM, 2012) is an AR application that allows users to see internal organs in a monitor and interact with them using natural gestures as illustrated in Figure 9.

Figure 9: By using mirracle, users are able to see internal organs overlaid on them through a monitor

Therefore, it represents one of the most cost efficient AR approaches since it requires off-the-shelf hardware components and standard PC equipment.

Some of its disadvantages are: small size of monitors which reduce immersion degree and FOV. Limited resolution of merged images and indirect remote interaction techniques (most commonly supported rather than direct interaction with real and graphic content) are also points of concern.

2.3.4. Spatial Displays

In contrast to body-attached displays, spatial displays seek to eliminate the need of using equipment attached to user’s bodies and, therefore, provide minimal intrusiveness. Spatial displays project virtual content directly on site. It allows multiple users at a time and, hence, enables collaboration.

They make use of video-projectors elements, holograms, radio frequency tags and tracking technologies to display graphic information directly in the real environment.

One example of this type of display is the projection-based which is a very promising technology according to (TORI, 2010). These displays use advanced computer graphics techniques to project graphic information on different physical objects surface, as exemplified in Figure 10 where children can play a soccer match on the floor.

Figure 10: Children can play a soccer match on the floor where the graphic data is projected

Projectors are becoming cheaper and smaller due to the advance of new technologies such as LED (light-emitting diode) and handheld projectors. However, mobility is still a problem for those displays since the setup for most projection-based displays is fixed (ZHOU, DUH and BILLINGHURST, 2008). Multiple projectors can be applied to increase the potential display area (BIMBER and RASKAR, 2006).

These displays offer improved ergonomics, a theoretically unlimited FOV, scalable resolution and easier eye-accommodation (BIMBER and RASKAR, 2006). Some of its disadvantages are: shadow casting of the physical objects and of interacting users and restrictions of the display areas.

2.4. Interaction Techniques

Interaction is one important aspect to take into account when one intends to design applications that are both appropriate and intuitive for its users.

(CARMIGNIANI, FURHT, *et al.*, 2011) state that AR interfaces are divided into four types: tangible AR interfaces, collaborative AR interfaces, hybrid AR interfaces and emerging multimodal interfaces. Their characteristics will be pointed out bellow.

2.4.1. Tangible AR Interfaces

In AR applications, there is an intimate relationship between virtual and physical objects. This suggests that one promising approach for good interface designs is to take advantage of the immediacy and familiarity of everyday physical objects for effective manipulation of virtual content. One way to do this is by using tangible AR interfaces (BILLINGHURST, GRASSET and LOOSER, 2005). These interfaces are powerful because they use the familiarity of everyday objects to ease the interaction.

In this type of interaction, object manipulations are mapped one-to–one to virtual object operations and follow a space multiplexed input design.

Some of its characteristics are: (i) each virtual object is registered to a physical object; (ii) the interaction with virtual objects is done through the manipulation of the corresponding physical objects. Thus, physical objects and interactions are as important as the virtual imagery and provide a very intuitive way to interact with the AR interface (BILLINGHURST, GRASSET and LOOSER, 2005).

One example of tangible interface is the AR Coloring Book (CLARK, DÜNSER and GRASSET, 2011). In this work, AR is used to enhance physical coloring books. Users are able to color the pages and the system recognizes them and produces 3D scenes

and textured models reflecting artwork created as seen in Figure 11.

Figure 11: Screenshot from the coloring book application

2.4.2. Collaborative AR Interfaces

Collaborative interfaces include the use of multiple displays to support remote and co-located activities (CARMIGNIANI, FURHT, *et al.*, 2011).

AR techniques can be used to allow users to move smoothly between virtual and real worlds. It also enables multiple users to be immersed on a scene and view augmented content in the setting. Therefore, it allows both face-to-face and remote collaboration by integrating multiple user's devices in different contexts and, therefore, enhancing telepresence.

(BILLINGHURST and KATO, 2002) list some of AR characteristics that facilitate collaboration: its ability to enhance reality; the presence of spatial cues for both face-to-face and remote collaboration; support for tangible interaction metaphor and ability to transition smoothly between reality and virtuality.

An example of a collaborative AR interface is the Construct3 (KAUFMANN, SCHMALSTIEG and WAGNER, 2000), a 3D geometric construction tool based on the Studierstube library. The system aims to foster mathematics and geometry education.

2.4.3. Hybrid AR Interfaces

Hybrid AR interfaces combine different but complimentary interfaces and displays (e.g.: opaque, stationary displays and see-through, head-worn displays) as well as the possibility to interact through varied interaction devices (CARMIGNIANI, FURHT, *et al.*, 2011).

In unplanned, everyday interactions, developers might not know beforehand what exact displays and devices will be used and who would be involved. These aspects might also change during the course of interaction. Therefore, hybrid AR interfaces provide a flexible platform which is able to accommodate a changing set of input and output devices as well as the interaction techniques.

One example of hybrid-tracking AR is the system developed by (SANDOR, 2005) which combines head-tracked, see-through and head-worn display to overlay AR and provides visual and auditory feedback. This system supports end users in assigning physical interaction devices to operations as well as virtual objects to perform them and reconfiguring the mappings between devices, objects and operations as the user interacts with the system.

2.4.4. Emerging Multimodal Interfaces

Emerging multimodal interfaces combine real objects input with natural forms of language (e.g.: speech, touch, natural hand gestures, or gaze). One example of this interface is the MIT's sixth sense (MISTRY P, 2009), or simply, WUW, which is a wearable gestural interface that allows the user to interact with the physical object through natural hand gestures, arms movement, and/or interaction with the object itself.

(CARMIGNIANI, FURHT, *et al.*, 2011) explain that this type of interaction has been largely developed and is heralded to be one of the preferred type of interaction for future AR applications as they offer a robust, efficient, expressive, and highly mobile form of human-computer interaction. Multimodal interfaces also have the ability to support users' ability to combine different modalities or to switch input modes depending on the setting or the task at hand.

2.5. AR IN EDUCATION

Since AR technology was first created it has been increasingly changing. In the early years, this type of technology was seen as bulky and heavyweight and appropriate mostly for industrial and military fields. Nowadays, advancements in different areas such as hardware and software and reduced costs enabled AR technology to spread through different fields.

(CARMIGNIANI, FURHT, *et al.*, 2011) point out that AR can be used for a range of fields such as professional training, medical visualization, entertainment purposes, advertising, maintenance and repair, robot path planning and many other areas.

Although all fields of knowledge can potentially take advantage from AR technology, (TORI, KIRNER and SISCOUTO, 2006) argue that teaching, learning and training will be particularly modified by the introduction of this new piece of technology and changes in interaction between teachers and students as well as students and information allowed by the mix of virtual and real information.

Those educational possibilities have been increasingly recognized by researchers. AR has also been considered one of the key emerging technologies for education over the next five years (JOHNSON, 2010).Coexistence of virtual and real information allows learners to visualize complex spatial relationships and abstract concepts such as in (SHELTON and HEDLEY, 2002). In this application, students are able to explore earth-sun relationships and, therefore, enhance their understanding of spatial and geographical concepts.

A classical example of AR applications for education is the MagicBook (BILLINGHURST, KATO and POUPYREV, 2001) which consists of a handheld AR display, a computer graphics workstation, and the physical book. By using the MagicBook, students are able to interact with the physical object (the book) as well as with AR since 3D content is projected on the book's pages. In addition, this book can also be used as an immersive virtual space. Students are immersed in the virtual space by seeing each other represented by avatars in the story setting.

More and more hardware and software can be used to create AR applications. An example of a recent head-attached device that can be used in AR applications is the Google Glass (GOOGLE). The Glassist (SILVA, FREITAS, *et al.*, 2014) is an AR application that uses this device and is meant to support teachers errands in a class, such as creating student's portfolios, managing their information and sharing them with peers.

As an example of tangible AR, we have the Augmented Chemistry (ALMGREN, CARLSSON, *et al.*, 2005), an application meant to scaffold organic chemistry learning. The system is composed by a booklet, a gripper, a cube, a platform, a camera and a

software. By using the gripper, the user can pick up elements from the booklet and add them to the molecule in construction on the platform as illustrated in Figure 12.

Figure 12: Augmented Chemistry in use

Another example is the AR Jigsaw Puzzle (SILVA, VILAR, *et al.*, 2014), an application that uses real puzzle pieces to display augmented content and, therefore, foster geographic skills. In Figure 13 is possible to see the AR Jigsaw Puzzle being used to complete a map of Brazilian states that received stadiums for the Fifa World Cup 2014.

Figure 13: AR Jigsaw Puzzle in use. It gives feedback when pieces are correctly put together

(WU, LEE, *et al.*, 2013) explain that advancements in handheld computing open up new opportunities for AR applications and create a subset of AR: mobile-AR.

The Cyberchase Shape Quest application is an example of a mobile–AR application that provides educational puzzles involving geometry, spatial reasoning and problem solving skills (PBS KIDS). Figure 14 presents a screenshot of the game.

Figure 14: Screenshot of Cyberchase Shape Quest game

AR has some affordances that may help to support learning. Some of them are: its ability to encourage kinaesthetic learning and the possibility to see virtual content in a 3D perspective. The coexistence of virtual and real environments also allows learners to experience phenomena that otherwise would be impossible in the real world, interact with two- or three-dimensional synthetic objects in the mixed reality environment and develop important practices and literacies that cannot be developed in other technology learning environments (WU, LEE, *et al.*, 2013).

(BILLINGHURST, 2002) adds that educational experiences offered by AR technologies are different from a number of reasons including: support of seamless interaction between real and virtual environments, use of tangible interface metaphor and ability to transition smoothly between reality and virtuality.

(SQUIRE and KLOPFER, 2007) present an AR game that can stimulate student's

prior knowledge and increase their academic engagement.

AR can also enhance collaboration between student-student and student-teacher. Additionally, many studies have shown that AR has a positive impact on student's motivation (DI SERIO, MARÍA IBÁÑEZ and KLOOS, 2013) (RADU, 2014).

AR can be used to leverage learning of different contents from math and science to human and arts. It is also possible to find AR applications for varied age levels ranging from young children to university students (L. CHOICE) (CAI, WANG and CHIANG, 2014) (REDONDO, FONSECA, *et al.*, 2013).

While AR offers new learning opportunities, it also creates new challenges for education in different domains, such as technological, learning and pedagogical issues. Technological issues may be related to cumbersome and expensive equipment as well as discomfort and poor depth perception (KERAWALLA, LUCKIN, *et al.*, 2006).

Learning issues may include cognitive overload by the large amount of information presented in AR applications, requirement to apply and synthesize multiple complex skills in spatial navigation, collaboration, problem solving, technology manipulation, and mathematical estimation and, finally, blend of reality and fantasy which can confuse students (WU, LEE, *et al.*, 2013).

Pedagogical issues, on the other hand, refer to instructional design and flexibility of the content. (WU, LEE, *et al.*, 2013) also refer as pedagogical issue different constraints from schools and resistances among teachers AR applications may encounter in a school setting.

Therefore, it is important to mind the gap between teaching and learning methods currently in use in most classrooms and the student-centered and exploratory nature of learning proposed by AR systems. Proper evaluation of AR tools is also required in order to understand its impact in the learning setting.

3. TEACHING ENGLISH TO YOUNG CHILDREN

Figure 15: Young children interacting with peers in a learning moment

Popular tradition encourages people to believe that young learners are better learners of a second language and far superior to adults.

Contrary to popular belief, research shows that children's success relies on a great deal of both cognitive and effective effort to internalize both native and second language.

A reasonable understanding of how children develop is a great tool for teachers and researchers. It helps them understand children better and be prepared to give them what they need. Many theories have been developed over time in order to explain and understand such complex phenomena. The progressive education movement, promoted by John Dewey, shared important ideas with Vygotsky, Montessori and Piaget. Its core ideas were that education should be child-centered, it should be both active and interactive as well as involve the social world of the child and the community (MOONEY, 2013).

3.1. THEORIES OF CHILDHOOD

Some representative authors and their most important thoughts are summarized below:

3.1.1. JOHN DEWEY

John Dewey was a defensor of a movement named progressive education which was a reaction to the rigid, formal style of traditional teaching during the nineteenth century. Some of the principles defended by Dewey were (MOONEY, 2013):

1. Education must emerge from the demands of the social situations the child finds himself in – In Dewey's opinion, children learn best when they interact cooperatively with peers and adults as illustrated in Figure 15 that shows children interacting with peers to read a book;
2. Children's own instincts and powers furnish the material and give the starting point for all education – That is, children's interests must be the basis of the curriculum. Teachers must consider children's interests and background when planning learning experiences;
3. Education is a process of living not a preparation for future living – He believed that education was part of life and therefore it must address what the person needs at the time not just prepare for the future. He argues that curriculum must be based on real work, home and social situations;
4. It is the business of the school to deepen and extend children's sense of value bound up in his home life – He believed that teachers must be sensitive to the family's values and cultures. There must be a connection between these values and what happens in schools;
5. Teachers are engaged not simply in the training of individuals but in the formation of proper social life – That means, they do not just teach school subjects but how to live in society. Therefore, they do not just teach individuals they shape society.

Dewey certainly believed that when children were engaged, learning was fun and exciting in and of itself. He argues that teachers must (MOONEY, 2013):

- Have great general knowledge as well as knowledge of specific children;
- Be willing to make sense of the world for children on the basis of their knowledge and experience;
- Invest in observation, planning, organization and documentation.

3.1.2. MARIA MONTESSORI

Although nowadays we take most of Montessori's ideas for granted at the time she opened her first school in 1907, they were a step ahead of the time. One of her premises is that children need child-centered environments to learn properly. She argued that appropriate size furnishing was really important, so most of the furnishing of her school she made by herself.

Some of the principles proposed by Montessori were (MOONEY, 2013):

1. Teachers must provide real working tools – She believed it was part of a sensory experience to have tools and utensils that fit children's hand and furniture that matches their bodies;
2. Children must work independently – This is required in order to children develop their sense of competence and responsibility. Montessori believed that adults spend much time serving children. She warned that children do not learn if they do not do things. She stated that children must be allowed to do everything they are able to do, this will help them feel competent and take responsibility;
3. The environment must be beautiful and orderly – Therefore, children can learn order from them;

4. Children learn through sensory experiences – According to Montesorri's ideas it is teacher's responsibility to provide different sights, textures, sounds and smells for learners;
5. Children need large blocks of open-ended time – Montessori's observation led her to believe that children are able of great concentration when they are surrounded by interesting things to do and are free to do them;
6. Teachers need to teach little and observe much – Montessori was formerly a scientist, so this influenced her to bring scientific skills to the classroom. She believed that if one is going to teach, one must know everything he/she can about those he/she hopes to teach. According to her ideas, the way to get to know children is to watch them.

3.1.3. ERIK ERIKSON

The works of Erik Erikson show how children develop the foundation for emotional and social development and mental health. He believed that in the earliest years of life, patterns develop that regulate, or at least influence, a person's actions and interactions for the rest of his/her life. He identified eight separate stages across lifespan. Each stage faces a crisis that needs to be resolved in order for people to develop socially and emotionally. (MOONEY, 2013) explains that the outcome of each stage is determined by the environment and caregiving strategies children are exposed to. However, it is always possible to go back and renegotiate issues from previous development stages. These stages can be seen in Table 1 adapted from (MOONEY, 2013).

Erikson's Stages of Psychosocial Development		
Age	**Stage**	**Strength Developed**
Birth to 12 months	Trust vs. Mistrust	Hope
1-3 years	Autonomy vs. Shame and Doubt	Willpower
3-6 years	Initiative vs. Guilt	Purpose
6-11 years	Industry vs. Inferiority	Competence
Adolescence	Identity vs. Role Confusion	Fidelity
Young adulthood	Intimacy vs. Isolation	Love
Middle age	Generavity vs. Self-Absorption	Care
Old age	Integrity vs. Despair	Wisdom

Table 1: Erikson's stages of psychosocial development

Since our concern is young children, we will discuss some characteristics of the first four stages and their implications for teachers.

3.1.3.1. TRUST X MISTRUST

Babies during this stage are developing a sense of trust in themselves, in other people and in the world around them. Erikson described trust as having two parts: external (belief that adults will meet their needs) and internal (belief in his/her own power to affect change and cope with diverse circumstances). The engagement of adults is part of what is called attachment. Children who lack strong attachments with adults

usually fail to develop empathy (MOONEY, 2013).

Teachers willing to help children develop proper trust skills can (MOONEY, 2013):

- Hold babies during feeding;
- Respond to signals of distress;
- Support babies attachment through primary caregiving.

3.1.3.2. Autonomy x Shame and Doubt

The aim of this stage is to acquire a sense of autonomy without suffering much of shame or doubt. Erikson points out that one of the main barriers for toddlers is overcontrolling behavior of adults.

Some attitudes teachers can have to help toddlers accomplish such a task are the following (MOONEY, 2013):

- Give children simple choices, such as: choose between limited options what they are going to wear, or what activity they are going to do first;
- Eliminate false choices – Do not give a choice to young children when there isn't really one, avoid rhetorical questions;
- Set clear limits for children – Teachers must set clear and firm limits for children;
- Accept alternating needs for independence and dependence – Teachers acceptance of this changing behavior helps toddlers grow in confidence and self-esteem.

3.1.3.3. Initiative x Guilt

The main task in this stage is to acquire a sense of purpose. According to Erikson (ERIKSON, (1950)1963), it is a time when children's development can go in two possible directions: human potential for glory or destruction. If pre-school children are encouraged to use their energy actively, their confidence will grow. On the other hand, if adults do for them what they already can do for themselves and focus on children's mistakes, their sense of initiative can turn into guilt and discouragement. Teachers of children at that stage can:

- Encourage children to be as independent as possible;
- Focus on gains as children practice new skills, not on the mistakes they make along the way;
- Set expectations that are in line with children's individual abilities;
- Focus curriculum on real things and on doing.

3.1.3.4. Industry x Inferiority

Children at this age range, 6 to 11 years old are in the early school years and, they will be learning how to read, write, do sums and do things by their own and, thus, developing a sense of industry.

Teachers begin to have an important role in the children's lives since they will be teaching them specific skills. Peer group also gain greater importance and is a source of self-esteem. The child needs to win approval by demonstrating specific competencies that are socially valued.

If children's potential are encouraged and reinforced, they develop a sense of industry and competence. However, if these stimuli is not provided and their initiative is not encouraged, children will begin to develop a sense of inferiority and doubt their own abilities. Therefore, they may not develop their full potential.

Ideally, elementary schools need to provide opportunities for students being valued and recognized by teachers, parents and their peers.

3.1.4. PIAGET

Instead of agreeing with current theories of his time that learning was either intrinsic (coming from the child) or extrinsic (imposed by the environment), Piaget (PIAGET, 1973) supported that children's interactions with the environment are what creates learning.

According to him, children's curiosity is what drives their learning. Piaget's theory invites teachers to be someone who nurtures inquiry and supports the children's own search for answers. He also highlighted the importance of play to learning.

Piaget's theory states that children's intellectual growth is based partly on physical development. Therefore, it is affected by children's interactions with the world. According to his ideas, teachers do not teach children a concept. They build their understanding through the things they do.

He divided learning in four stages as can be seen in Table 2:

Piaget's Stages of Cognitive Development		
Age	**Stage**	**Behaviors**
Birth to age 2	Sensorimotor	Learn through the senses; learn through reflexes; manipulate materials.
2-7 years	Preoperational	Form ideas based on their perceptions; can only focus on one variable at a time; overgeneralize based on limited experience.
7-11 or 12 years	Concrete Operational	Form ideas based on reasoning; limit thinking to objects and familiar events.
11 or 12 years and older	Formal Operational	Think conceptually; think hypothetically.

Table 2: Piaget's stages of cognitive development

In the sensorimotor stage, babies rely on their senses and physical activity to learn about the world and develop their cognitive understanding.

In the preoperational stage, children's thinking differs from adult patterns. They are egocentric (think of everything only for their own point of view). In addition, they do not fully understand conservation properties and tend to believe in what they see. Therefore, they do not have a firm grasp of qualities belonging to objects in their world.

As children go to the concrete operational stage, there is a change in thought pattern which is characterized for its reversibility (children are able to reverse the direction of their thought). Children in this stage, for example, no longer count on their fingers. They begin to notice differences in classes of objects.

Finally in formal operations stage, they develop the ability to think logically and in hypothetical terms.

3.1.5. VYGOTSKY

According to (MOONEY, 2013), Vygotsky's theories were – and continue to be – controversial. He objected the analysis of children based on intelligence tests. He believed that careful observation (qualitative

research) of children should be considered as valid scores on a test (quantitative research).

Vygotsky thought that personal and social experiences couldn't be separated. Like Piaget, he argued that much learning takes place during play. Language and development build on each other. When children play they constantly use language, discuss roles and determine conditions to make-believe. These interactions contribute to children's knowledge development.

Vygotsky's theory establishes an important concept – the zone of proximal development (ZPD), which is defined as the distance between the most difficult task a child can do alone and the most difficult task a child can do with help. A child on the edge of learning can benefit greatly from interaction with teachers and their peers. This assistance is called scaffolding.

Similarly to John Dewey's belief, Vygotsky understands that teachers must use their knowledge of the world to help children make sense of it. He also emphasizes the importance of observation. According to him, this is the only way teachers can accurately assess what is a child's ZPD at a given moment. This knowledge is crucial to plan curriculum according to children's needs.

Vygotsky helped teachers to see that children not only learn by doing, but also by talking, working with peers and persisting in a task. He, therefore, emphasizes the importance of language and interaction among children (MOONEY, 2013).

There is growing evidence that a pre-school child's ability to apply cognitive control, also named, executive function[1], is a better predictor of later school success than any academic learning. Vygotsky's concepts of ZDP and scaffolding are thought to be helpful to foster self-regulation skills (MOONEY, 2013).

3.2. CHARACTERISTICS OF YOUNG CHILDREN AND THEIR IMPLICATIONS FOR ENGLISH TEACHING

Overall, children and adults vary in their abilities and needs. According to (BROWN, 2007) the difference between children and adults learners is primarily the contrast between child's spontaneous, peripheral attention and language forms and the adult's overt, focal awareness of and attention to those forms.

Studies have shown that adults can be superior in a number of aspects in acquisition of language, such as, vocabulary retention and grammatical learning. In addition, children's fluency and naturalness can be hampered depending on the context of classroom instruction.

Teaching English as a second language to school-age children is not merely a matter of setting them loose on a plethora of authentic language tasks in the classroom. Successful teaching requires specific skills and intuition that differ from those appropriate for adult teaching (BROWN, 2007).

It is important to consider that there is a big difference between what very young children (four to six years old), pre-pubescent children (twelve to thirteen years old) and the

[1] According to [54], executive functions refer to the ability to manage basic and cognitive and emotional processes, such as self-regulation, the ability to focus on tasks, the ability to organize thoughts and materials, and the ability to go through and complete tasks. Emotional processes, such as self-regulation, the ability to focus on tasks, the ability to organize thoughts and materials, and the ability to go through and complete tasks.

whole range of ages in between can do. Children develop differently. Some children develop early, some later; some gradually, others in leaps and bounds. (SCOTT and YTREBERG, 1990) point out that although it is not possible to say that all children of five can do *x*, it is possible to highlight some characteristics of young children which teachers should be aware of and take into consideration while teaching.

Regarding children's development, there are five categories that may help to give some practical approaches to teaching children. These characteristics are as follows.

3.2.1. Intellectual Development

Children (up to the age of about eleven) are in the 'concrete operations' stage as defined by (PIAGET, 1972). As pointed out earlier in this chapter, this means that they still have some limitations concerning rules understanding. According to (HARMER, 2001), they respond to meaning even if they do not understand specific words. They often learn indirectly rather than directly which means they take information from all sides. Children learn from everything around them.

Children's understanding comes not just from information, but also from what they see and hear, and, especially, have a chance to interact with. This implies that rules stated in abstract terms should be avoided.

(VALE and FEUNTEUN, 1995) explain that children may not see any purpose in learning a given grammar structure just because it is in the next page in the coursebook. The authors add that children may learn a grammar structure in a coursebook context and make no connection with its use in their terms. Until children can see how a particular point affects their world, the learning of a structure may be purely by rote, and easily forgotten (VALE and FEUNTEUN, 1995) Hence, it makes sense to provide the purpose (e.g. a practical task) first and the tools for describing it later (e.g. the present simple tense).

3.2.2. Attention Span

One important difference between young and adult learners is attention span. Children are known to have a limited attention span, unless tasks are extremely engaging.

In the words of (SCOTT and YTREBERG, 1990), short attention span comes into play when children have to deal with activities that are boring, useless or too difficult. In addition, the authors explain that since language lessons can be at times difficult for children, teacher's job is to make them interesting, lively and fun.

These authors go further and suggest some alternatives for teachers on how to have more engaging and productive classes. These suggestions are described below:

- Design activities that capture children's immediate interests;
- Incorporate varied activities in the lesson. This suggestion is also reinforced by different authors (VALE and FEUNTEUN, 1995);
- Being animated, lively and enthusiastic about subject matter;
- Have sense of humor;
- Take advantage of children's natural curiosity about the world (HARMER, 2001).

3.2.3. Sensory Input

Studies suggest that children need to have all five senses stimulated. Therefore, teachers should strive to go beyond visual

and auditory modes that are usually the most stimulated in the classrooms.

This can be achieved by a plethora of tasks, such as, physical activities, games, roleplaying, hands on activities, sensory aids (e.g. the smell of flowers or taste of food). Nonverbal language should also be stimulated.

3.2.4. Affective Factors

Although children are often creative in language forms that does not mean they do not have inhibitions. Young learners are extremely sensitive especially to peer pressure. According to (SCOTT and YTREBERG, 1990), children are more fragile than adults in many ways. Their egos are still being shaped, hence, slightest nuances of communication can be negatively interpreted.

(VALE and FEUNTEUN, 1995) state that in any learning situation, where individuals are required to act and interact with others, there are many potential social and emotional (affective) constraints and pressures that may interfere with effective learning.

Teachers need to help students overcome such constraints. The effect of lowering these learning barriers facilitates learning to take place. It also encourages a wider sharing and exchange of ideas. The content studied becomes of a higher quality and greater quantity (VALE and FEUNTEUN, 1995).

3.2.5. Authentic, Meaningful Language

According to (BROWN, 2007), children are focused on what a new language can be actually used for. The authors highlight that children have a good sense of what is not authentic. Therefore, language needs to be context-embedded. A whole language approach is essential. If language is broken into too many bits and pieces, children will not be able to see the relationship to the whole and will not stress the interrelationships among the various skills (listening, speaking, reading and writing).

In order to teach the four language skills to young children, it is important to keep in mind their different abilities and needs. Young children's class is usually different from adult's classrooms. Children are usually in movement (making things, holding things, moving their hands or working somewhere). Young classes sometimes look like an art class (PECK, 2001).

3.3. Teaching Listening and Speaking to Young Children

(PECK, 2001) points out some important contrasts between children and adult learners:

- Children are more likely to play with language than adults;
- Children can become more engaged through stories and games;
- Younger children are less likely to notice errors or correct them;
- In general, children are more holistic learners who need to use language for authentic communication in ESL [2] (English as Second Language) classes.

Therefore, many researchers highlight the need to teach children holistically. It is important to provide varied material. The activities must be child-centered and

[2] Authors usually refer to ESL as English learned within a culture where English is natively spoken [106].

communication authentic. Some common themes regarding teaching children are the following (PECK, 2001):

- Focus on meaning, not correctness;
- Provide a rich context, including movement, the senses, objects and pictures, and a variety of activities;
- Teach ESL holistically, integrating the four skills;
- Treat learners appropriately in light of their age and interests;
- Treat language as a tool for children to use for their own social and academic ends;
- Use language for authentic communication, not as an object of analysis.

Since the 1960s, oral language has been emphasized more than written language in children's ESL learning (PECK, 2001). In some ways, children approach oral language differently than adults do. For instance, children appear more likely to play with language (PECK, 1978), they may learn through language play (PECK, 1980), enjoy rhythmic and repetitive language more than adults, play with intonation of a sentence and are more willing to sing, enjoy repeating a word or an utterance in a play situation, are less aware of the ways in which languages can differ, are more likely to laugh at the sounds of a second language or remind a word in the first language, young children may comfortably talk to themselves perhaps as part of a fantasy role play.

It is recommended for teachers to use a variety of materials and activities in order to explore students characteristics (REILY and WARD, 1997) (LEWIS, 2004). They could explore songs, poems or chants as well as dramatic activities and internet resources.

Another feature that is important to explore is gesture and movement. Children need to move around more than adults do. In regard to this, there is a teaching approach named Total Physical Response (TPR) that involves movement (LARSEN-FREEMAN, 2000). In this approach, teachers give commands (e.g. stand up, touch your shoulders), model them and gradually lead students to follow the commands.

Grammar is another feature of language that must be taught, however, young children are less likely to focus on the vocabulary or pronunciation errors of others, or to correct them. As children grow older, their metalinguistic awareness (ability to analyze language) grows and they tend to notice errors much the same as adults do.

3.4. Teaching Children Literacy Skills in a Second Language

(EDIGER, 2001) points out that in recent years there has been an increased focus on the teaching of reading and other literacy skills to children both in America and abroad. She explains that such an interest may be due to some factors, for instance, that many states and provinces in the US and Canada have established literacy tests and standards and that a growing number of countries are making English language instruction mandatory. It is important to understand, however, the dimension and complexity of such skills in order to teach them effectively. Literacy skills are not just an additional skill one needs to learn in schools, but, an integral part of one's lives (CAMERON, 2001).

Literacy[3] is both a social and cognitive skill. It provides people with opportunities to share meanings across space and time (social dimension) and it involves the use of specific skills and knowledge about how written language operates in texts (cognitive dimension) (CAMERON, 2001).

Reading is an interactive, sociocognitive process involving a text, a reader and social context (EDIGER, 2001).

Cognitive and language processes of literacy involve a plethora of skills and knowledge at different levels. Readers and writers need to develop different skills, namely (GRABE, 1991):

1. Automatic recognition skills – recognize text, especially words;
2. Vocabulary and structural knowledge – understanding of language structure (e.g.: how syllables make up words, word order) and large vocabulary recognition;
3. Formal discourse structure knowledge – understanding of how texts are organized and knowledge of different textual genres (e.g.: poems, letters, recipes);
4. Content/World background – understanding of world information related to the text at hand;
5. Synthesis and evaluation skills – ability to read and compare information from varied sources. Ability to read critically and select relevant and/or useful information;
6. Metacognitive knowledge and skills monitoring – awareness of mental processes involved in the reading task and strategies employed while reading.

Skilled readers are able to use all those strategies listed above. This is a particularly difficult task for young readers especially when learning a foreign language. Some important aspects concerning these issues are the following (EDIGER, 2001):

- They usually have little or no oral knowledge to rely on, thus, language and vocabulary in reading is often completely new to them;
- Aspects of oral ability are still developing;
- Oral language proficiency is not necessarily an indicator of reading and writing abilities;
- Children do not need to be orally fluent to start learning how to read and write. For instance, studies have shown that they are able to begin reading what they see in the environment around them (e.g.: identify meanings of words of packaged products or signs);
- There is a strong relationship between children's prior native language literacy and English literacy development;
- Relationships among listening, speaking, reading and writing during development are complex relationships of mutual support.

(CAMERON, 2001) explains that since the mid-1980s the foreign reading learning debate was dominated by two major approaches, the top down (emphasizing meaning making) and the bottom-up (emphasizing the knowledge of lower level skills and recognition of letter sounds links). Many approaches have been created between whole language and phonics approaches.

[3] The term literacy refers to the condition assumed by one who learns how to read and write [101].

(CAMERON, 2001) emphasizes that such oppositions are artificial and do not help students who need to master all those abilities.

In what concerns this chapter, we will list some aspects teachers might take into account while teaching young students based on current research.

As (EDIGER, 2001) suggests, teachers need to be prepared to employ a variety of teaching techniques and approaches with the students. They should also learn as much as possible about student's cultural background and experiences.

The author goes further and suggests some ideas for teachers to foster students' literacy skills:

- Provide opportunities for children to read more extensively on a subject – this can be accomplished through content study, investigation and internet research and projects;
- Provide authentic purposes for reading and writing – set up real communication contexts for student's activities;
- Provide scaffolding for learning – this suggestion is based on Vygotsky's concept of scaffolding. It has to do with giving the appropriate support students need to develop their abilities. As students become able to work with complex language, supports can be decreased or simply removed;
- Use oral skills to support reading and writing development – the author encourages the use of group discussion as a support for the development of reading and writing skills;
- Focus student's attention on reading and writing strategies – the author suggests teachers to call students attention to strategies employed while reading and monitor their use in order to ease understanding.

3.5. TEACHING 21ST CENTURY SKILLS AND COMPETENCES

Many educators agree that what we teach to young people, and the way we teach them, must change. (BAKER, 2010) explains that teachers can no longer afford to ignore the presence of new media, such as the internet, television, music, or movies. He emphasizes that these media are the domain of today's students, just as their mobile phones have become the convergent media tools of tomorrow.

Some authors even argue that if students aren't taught the language of sound and images, they should be considered as illiterate as if they left college without being able to read and write.

Media literacy is an important topic to be integrated throughout the curriculum so that students have an opportunity to become actively engaged in learning about it multiple times and in multiple ways.

Apart from the access provided at schools, young learners have access to different media sources beyond the school walls and therefore they need critical skills to properly deal with this huge amount of information.

While more young people have access to the internet and other media this does not necessarily mean they have the intellectual skills or the predisposition to critically analyze the content received and produced.

Therefore, students need skills to live and cope with real world issues. They need to master competencies for work, citizenship and life-long learning.

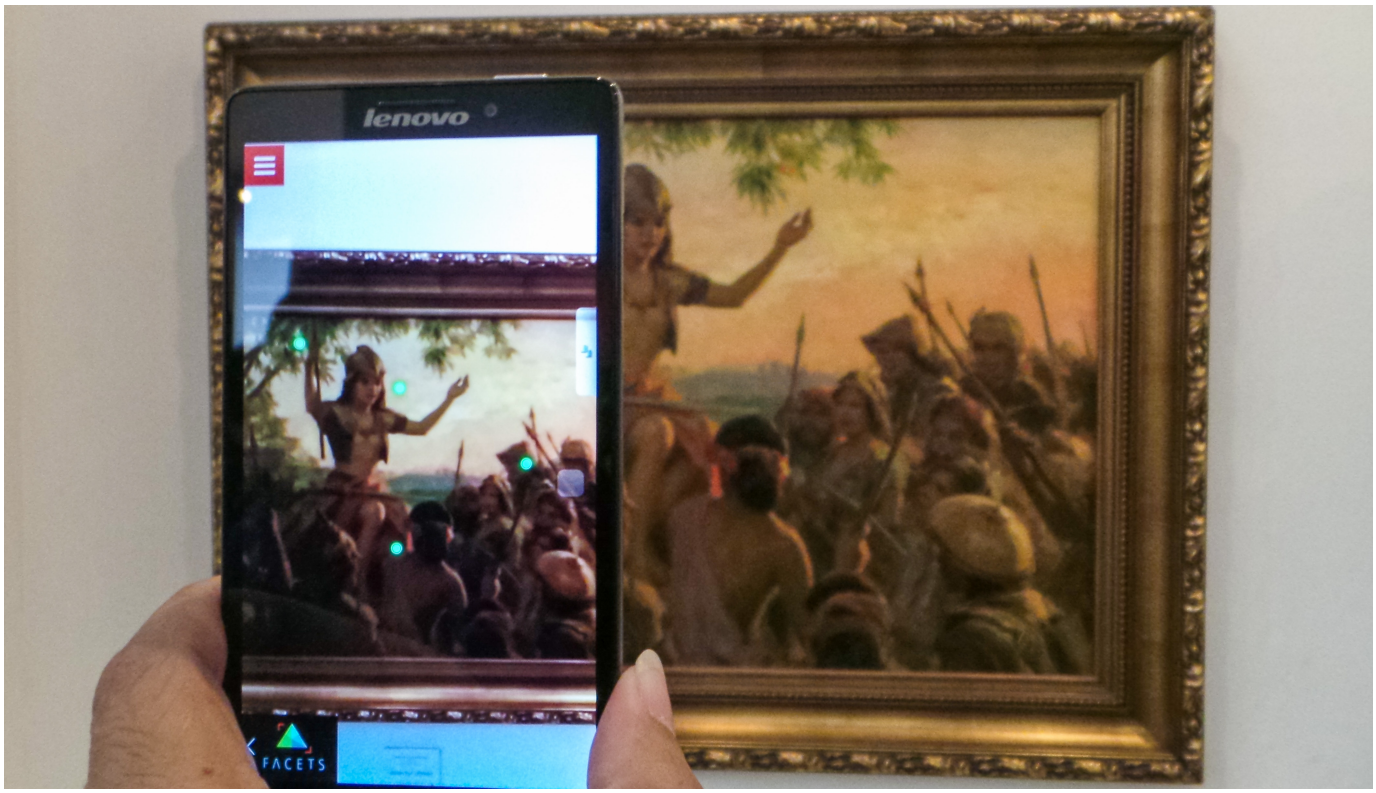
(WEST WINDSOR PLAINSBORO REGIONAL SCHOOL DISTRICT, 2009) recommends that 21st century education must be founded on solid academic knowledge and build upon this foundation with six competencies that will be described below:

- Collaborative team member – students need to work with and learn from diverse groups, be flexible and adaptable;
- Effective communicator – students need to learn how to express thoughts clearly, articulate opinions, motivate and communicate coherently. They must also learn to consider and understand how images and words are shaped by current technologies;
- Information literate researcher – students must be given the tools required to take advantage of available information. Information literacy is the basis for lifelong learning and enables them to master content and extend investigation and, thus, taking control of their own learning;
- Innovative and practical problem solver – students need to master a set of skills related to problem solving which is the process of applying scientific and engineering methods of defining and describing a problem, generate, implement, monitor and evaluate possible solutions;
- Global aware, active, and responsible student/citizen – students need genuine experiences living, studying and working within cultures different from their own;
- Self-directed learner – students need to become reflective about thinking and learning as well as develop strategies to process information effectively and be self-confident in their abilities to succeed.

Thus, students need those skills in order to cope with the world as it is today and have the necessary abilities to adapt and succeed in the future. Academic content is the means for developing competencies, instead of being the goal, as it has been traditionally (WAGNER, 2008).

4. EVALUATING AR IN EDUCATION

Figure 16 AR application developed to enhance museum experience



It has been long since AR's potential in education is being investigated. According to (KOSTARAS and XENOS, 2009), AR can aid learning and make the overall learning process much more interesting and pleasant. In a rapidly changing society as ours where there is a great amount of information available, it is of major importance to know how to locate information and use it efficiently for both education and business. AR is a technology that dramatically shifts the location and timing of education and training (LEE, 2012).

(BILLINGHURST and DUNSER, 2012) point out that unlike other computer interfaces that draw users away from the real world and onto the screen, AR interfaces enhance the real world experience as illustrated in Figure 16. This figure presents an application designed to create new museum experiences through the use of AR (AMOR, 2014). The authors also highlight some reasons why educational experiences afforded by AR are different: (1) Support of seamless interaction between real and virtual environments, (2) The use of a tangible interface metaphor for object manipulation and (3) The ability to transition smoothly between reality and virtuality.

Although AR has been studied for over forty years it has only been recently that researchers have begun to formally evaluate AR applications (DÜNSER, GRASSET and BILLINGHURST, 2008). The authors point out that one reason for the lack of user evaluations in AR may be, among other factors, a lack of knowledge on how to properly evaluate AR experiences and design experiments. According to them, there seems to be a lack of understanding regarding the need of doing studies and the right motivation for carrying them. If user evaluations are conducted out of incorrect motivation or if empirical methods are not properly applied, the reported results and findings are of limited value or can even be misleading.

So far the amount of AR systems formally evaluated is rather small (DÜNSER, GRASSET, *et al.*, 2007). For example, literature surveys of user evaluation in AR have found that only around 8% of published AR research papers include formal evaluations (SWAN and GABBARD, 2005) (DÜNSER, GRASSET and BILLINGHURST, 2008). According to (DÜNSER and BILLINGHURST, 2011), one reason for this small percentage may be the lack of suitable methods for evaluating AR interfaces.

Researchers in emerging interface fields such as Virtual Reality (VR) or AR cannot rely solely on design guidelines for traditional user interfaces since new interfaces afford new forms of interactions (DÜNSER and BILLINGHURST, 2011).

When it comes to applications devoted to learning, it is very important to evaluate their impact on the learning experience and the feasibility of incorporating them into the classrooms. There may be many factors involved varying from cost to staff's acceptance. Evaluation of technology is an important step in design instruction. Hence it is necessary to evaluate it properly so practitioners are more confident in its positive effects on the learning process. It is also relevant to consider the point of view of both teachers and learners since they might differ. For instance, (BALOG and PRIBEANU, 2010) had shown the perceived usefulness and the perceived enjoyment as relevant factors for student's acceptance of an AR application, while the perceived ease of use was not a significant factor for student's acceptance.

A recent survey reviewed applications intended to complement traditional curriculum materials for K-12 (SANTOS, CHEN, *et al.*, 2014). In this work the authors performed a qualitative analysis on the design aspects and evaluation for AR Learning Environments (ARLES). The focus of the survey was to investigate ARLES designed for kindergarten, primary and/or secondary school. Its aim was to explore learning theories as basis for effective learning experiences. They found out that there are three inherent affordances of AR to educational settings: real-world annotation, contextual visualization, and vision-haptic visualization. It was also noted that these affordances are supported by existing theories: multimedia learning theory, experiential learning theory and animate vision theory. Concerning the evaluation aspects, the authors discovered that aside from the performance of students in pre-tests and post-tests, other aspects of the learning experience such as motivation and satisfaction were usually observed in the evaluations performed in the literature.

4.1. RESEARCH QUESTIONS

Taking into account the complexity of the classroom environment and its implications in the process of technology acceptance and use, it was conducted a systematic review in order to investigate how researchers have been evaluating their AR systems. The main question of our review was “how researchers evaluate AR based educational technology?”. In order to provide an overview of research in this field and guide data extraction, analysis and synthesis, several subquestions were formulated as described below. The questions were classified into three categories: descriptive, classificatory, and relation and effect.

Descriptive questions:

1. What is the evolution in number and type of research from 2009 to July 2014?
2. What institutions are most involved in performing this type of research?

Classificatory:

3. What are the different designs (methodology) used in these studies?
4. What are the target populations used in these studies?
5. What are the constructs being analyzed?
6. What are the domains of the different applications tested?
7. What types of research questions are investigated?
8. What are the types of AR technology used?
9. What is the problem being analyzed?
10. What is the AR definition being used?
11. Is the application based on educational theory?
12. What technologies AR is combined with?
13. How was the involvement of teachers in the evaluation process?
14. Did the study use multiple metrics (both quantitative and qualitative)?
15. Did the study use multiple metrics for educational evaluation purposes?

Relation and effect:

16. What are the implications of the findings in research and practice?
17. What is the impact of the tool analyzed (positive or negative)?

4.2. METHODOLOGY OF THE SYSTEMATIC REVIEW

The literature review was performed following the PRISMA protocol (MOHER, LIBERATI, *et al.*, 2009). The first step of the review was to establish the search string for the paper selection. The search string was created based on the research questions previously mentioned. The terms were defined along with synonyms found in the body of literature as shown in Table 3.

Search String
(“augmented reality”) AND (“education” OR “learning”) AND (“evaluation”) AND (“educational tool” OR “educational system” OR “educational application” OR “educational platform”)

Table 3: Search strings used in the systematic review

Then, the databases for the systematic review were defined. The papers were searched automatically in three databases as described in Table 4.

Search Databases	
ACM	http://dl.acm.org
IEEE Xplore	http://ieeexplore.ieee.org/Xplore/home.jsp

Science Direct	http://www.sciencedirect.com

Table 4: Search databases used in the systematic review

For papers to be included in the study, they must meet the following criteria:

1. Papers published in the English Language (Papers in other languages were excluded. Dissertation, thesis, presentations, abstracts, technical reports and short-papers were also excluded). Short papers encompassed works up to 4 pages.
2. Papers were only considered one time (In case of repetitive papers, we considered the more complete one or the most recent).
3. Papers published from 2009 until July 2014.
4. Papers that explicitly mentioned their evaluation methodology.
5. The papers must have at least an AR prototype working.
6. The AR solution must be tested with its end users.
7. The solutions presented must be applied to learning a new concept or skill.
8. Papers that intended to evaluate learning aspects.

The search strings were used in the three selected databases. Firstly, a search was performed in the databases using the search strings. A total of 887 articles were found out after this step. After removing duplicates, 880 articles were identified.

Secondly, the researcher screened the papers by reading their title, abstract and conclusion in order to eliminate the ones that were clearly not related to the research question. This step was entitled pre-selection phase. 49 papers were left after this step.

The next step was to evaluate those papers applying the inclusion criteria previously mentioned. After this step, 14 papers were eligible for the study[4].

4.2.1. Quality Criteria Evaluation

A quality assessment of papers included in the review was performed both in terms of their quantitative and qualitative results.

The QualSyst standards developed by the Healthy Technology Assessment (HTA) research group was used as guideline for quality control (L. M. KMET, 2004).

The questionnaire proposed consisted of 14 items evaluating study questions concerning the following aspects: design methodology, sample, outcomes, results outcomes, description and conclusions. Some items such as evaluator and user blinding were not scored due to the non-applicability in the study's methodology. Other items such as interventional and random allocation varied from study to study being applied only in some cases.

According to the classification proposed, each item was graded as it fulfilled the requirements in three categories: total, partial and none with assigned scores of 2, 1 or 0 respectively. The total sum was divided by the maximal possible points (e.g. 10 items x 2 points = 20 points). The final score of each included review paper was presented as a grade. The results of this analysis will be properly presented in the quality of report subsection.

[4] In Appendix A, the list of the selected papers is presented.

4.3. DATA EXTRACTION

In this phase, we answered some questions from the selected papers analyzed. An excel spreadsheet was used to organize the data collected. The questions can be seen below:

- Title;
- Year;
- Authors;
- University/Research group;
- Source (Conference or journal);
- Methodology design;
- Target population;
- Application domain;
- Type of research question;
- Implications for practice;
- Type of AR technology (tracking, display, interactions);
- AR definition;
- What constructs does it evaluate?
- Is the application based on educational research?
- What technologies AR is combined with?
- How was the involvement of teachers in the evaluation process?
- Did the study use multiple metrics (both quantitative and qualitative)?
- Did the study use multiple metrics for educational evaluation purposes?
- What are the implications of the findings in research and practice?
- What is the impact of the tool analyzed (positive or negative)?
- Observations.

4.4. RESULTS AND DISCUSSION

In this section we will describe and discuss the results of the systematic review process by analyzing the quality of report of selected papers and how they answered our research questions.

4.4.1. QUALITY OF REPORT

As regards to the quantitative criteria, overall, studies presented good quality. The majority of papers received grades higher than 9[5].

However, S14 did not clearly describe some aspects of its methodology. It is important to highlight that not all the studies presented a control group to compare the results with (S01, S13, S14). When dealing with schools, sometimes is not possible to have a control group as reported in S01.

Depending on the nature of the studies, sometimes it was not possible to have randomized trials as in studies S01, S04, S08, S10, S14. The methodologies of these studies (mostly, surveys and semi-experiments) did not apply for this condition. Some papers did not mention random allocation (S02, S03, S09 and S12).

Some studies reported that their sample sizes were not enough for performing statistical analysis (S01, S04).

Through the quantitative analysis, it was noticeable that some confounding effects were difficult to control, such as history effects (student's background), memory decay and the novelty effect played by a new technology, such as AR. These aspects were mentioned in S01, S04, S08, S13 and S14.

As for qualitative aspects, the studies overall presented good quality of report. Almost all studies received full grade[6].

It is important to highlight though that not all of the studies presented qualitative measurements (S03, S05, S06, S11, S12,

[5] [5] In Appendix B, the quantitative criteria is detailled.
[6] [6] In Appendix C, the qualitative criteria is detailled.

S14). From the ones presenting qualitative measures, not all of them used those measurements to evaluate learning outcomes. For instance, studies S04, S07, S08 and S13 did use qualitative measures to evaluate learning outcomes. Indeed, S04 only used qualitative measures to evaluate this aspect.

S01 used qualitative aspects to perform a preliminary study in order to introduce their prototype in the school and evaluated it with a panel of teachers. Authors reported they identified some additional functionality incorporated in the final version of the tool. S02 also used qualitative measures (expert survey) to refine its tool.

S09 used qualitative standards to identify learner's knowledge construction behavior according to an adapted three category coding scheme. Finally, S10 used interview questionnaires to explore museum visitor's attitudes toward the use, acceptance and advantages and disadvantages of the guide systems.

4.4.2. Descriptive Questions

The questions in this category were: (1) what is the evolution in number and type of this category of research from 2009 to July 2014? and (2) what institutions are mostly involved in performing this type of research?

Through Figure 17 it is possible to visualize that the research in this field is steadily growing. However, it is noticeable that there were no research between 2010 and 2011.

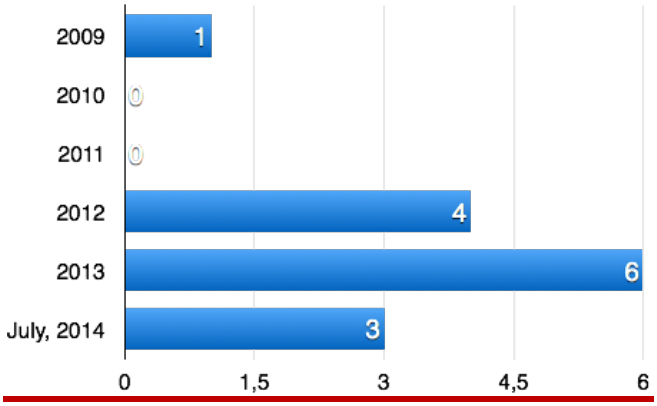


Figure 17: Papers according to the year of publication

Table 5 presents the institutions involved in the research.

Study ID	Institution
S01	Universidad CEU Cardinal Herrera
S02	Korea University
S03, S05	Universitat Politècnica de Catalunya
S04	University of Birmingham
S06	National Chengchi University
S07	Universitat Politècnica de València
S08	Harvard University
S09	National Taiwan University of Science and Technology
S10	National Taiwan Normal University
S11	Pontificia Universidad Católica de Chile
S12	Universidad Ramon Llull
S13	Beijing Normal University
S14	University of Helsinki

Table 5: Institutions involved in the research

It is noticeable that most of the institutions involved in the research are located in Europe (8), 5 of them are in Asia and 1 in Latin America.

4.4.3. Classificatory Questions

The questions in this category were: (3) what are the different designs (methodology) used in these studies?; (4) what are the

target populations used in these studies?; (5) what are the constructs being analyzed; (6) what are the domains of the different applications tested?, (7) what types of research questions are investigated?, (8) what are the types of AR technology used?; (9) what is the problem being analyzed?; (10) what is the AR definition being used?; (11) is the application based on educational theory?; (12) what technologies AR is combined with?; (13) how was the involvement of teachers in the evaluation process?; (14) did the study use multiple metrics (both quantitative and qualitative)? and (15) did the study use multiple metrics for educational evaluation purposes?

As regards to methodology design, Table 6 shows the methodology used in the selected studies.

Study ID	Research Design
S01	Prototype development; evaluation of the prototype with a panel of teachers; quasi-experimental design scheme based on interrupted time series; questionnaire
S02	Preliminary expert survey, followed by a system refinement according to the survey results, and an in-field learning effects assessment study (formative evaluation)
S03	Case study (project architecture development); questionnaires
S04	Experiment; pre-test; post-test; observations
S05	Pre-test; post-test; usability questionnaire
S06	Experiment; questionnaire; pre-test; post-test
S07	Questionnaires and observations
S08	Pre-field training trip; a field trip to a local pond environment; post-field discussion trip in the classroom; survey; interview
S09	Pre-test; post-test; tape-recording
S10	Quasi-experimental design; pre-test; post-test; interview
S11	Experiment; pre-test; post-test
S12	Pre-test, a technological profile test, a post-test usage; satisfaction test, a structured test was used with the intranet moodle system of the university
S13	Interview; pre-test; post-test; questionnaire; video-recordings (observations)
S14	Likert-scale questionnaire and three T-test tests, wilcoxon and marker test

Table 6: Study design of selected works

Through this table, it is possible to observe that most of the studies combined different methodologies in order to evaluate their tools.

The majority of papers used a pre-post test design (S04, S05, S06, S08, S09, S12 and S13). S12 combined this design with a technological profile test and a satisfaction test. Quasi-experimental designs were used in three studies S01, S02 and S09.

According to (COHEN and MANION, 1989), the essential feature of experimental research is that the investigator deliberately controls and manipulates the conditions, which determine the events in which he is interested in. (EASTERBROOK, SINGER, *et al.*, 2008) points out that a precondition for conducting an experiment is a clear hypothesis since it will guide all the steps of the experimental design, including deciding which variables to include in the study and how to measure them. These authors also

highlight that variants on experiments are possible and can be used whenever a true experiment is not possible. For instance, quasi-experiments may be used when subjects must be allowed to choose their treatment. On the hand, in time-series experiments, the effect of a treatment is measured in discrete time steps over a period of time.

A large amount of papers used questionnaires as a metric for evaluation (S01, S02, S03, S05, S6, S13 and S14).

Questionnaire is a research instrument that consists in a series of questions or prompts aimed at gathering information from subjects. According to (COHEN and MANION, 1989) an ideal questionnaire has the same properties as good law, that is, it must be clear and unambiguous. Its design must avoid potential errors from respondents.

A prototype evaluation by experts was also done in two studies (S01 and S02). Prototype evaluation by experts was used by S01. The main goal of expert reviews is to identify potential usability problems, and check conformity with usability principles, namely, effectiveness, efficiency, engagement, error tolerance, and ease of leaning (QUEENSBERRY, 2011).

Interview was used by S08, S10 and S13. Interview is a process in which a researcher and participant engage in a conversation focused on questions related to a research study (DEMARRAIS, 2004). (MERRIAM, 2009) argues that interviews are used to find out things that cannot be directly observed, such as feelings, thoughts and intentions. Interviews can vary in structure from the highly structured to the unstructured ones. However, the semi-structured is most commonly used. Interviews can be in groups. In this case, they are named focus groups interviews. Some key aspects in interviews are asking good questions and purposeful sample. The author also points out the importance of considering how to begin the interview and account for the complexity of the relationship between interviewer and interviewee, which will result in a more informed analysis of the data collected.

Case study was used by S03. Case study is an empirical inquiry that investigates a contemporary phenomenon within its real-life context, especially when the boundaries between phenomenon and context are not clearly evident (YIN, 2008).

(MERRIAM, 2009) points out that the most defining characteristic of a case study lies in delimit the case to be studied. Thus, case study research uses purposive sampling rather than random sampling (EASTERBROOK, SINGER, *et al.*, 2008). Researchers usually observe characteristics of an individual unit (a child, a class, a school or a community). A variety of resources are typically used to collect data and qualitative data plays a central role as they tend to offer rich insight into the case. This type of survey is suited when it is impossible to separate the phenomenon variables from their context (YIN, 2008). According to (EASTERBROOK, SINGER, *et al.*, 2008), the major weakness of this type of study is that the data collection and analysis is more open to interpretation and research bias, thus, authors argue that it is necessary to follow an explicit framework to select cases and collect data.

Surveys were used by S02 and S08. Surveys are perhaps the most commonly used descriptive method in educational research. It is used to identify the

characteristics of a broad population of individuals. According to (COHEN and MANION, 1989), surveys usually gather data with the following intentions: (a) describe the nature of existing conditions, or (b) identify standards against which existing conditions can be compared, or (c) determine the relationships that exist between specific event. This data collection can be done through the use of questionnaires, structured interviews, or data logging techniques [ref]. Some of the issues involved in this type of research are the resources available, the selection of a representative sample from a well-defined population and issues related to validity that can be seen from two points of view: (a) whether respondents who completed the questionnaires do so accurately and (b) whether those who fail to return their questionnaires would have given the same distribution of answers as did the returnees (COHEN and MANION, 1989).

Thus, (EASTERBROOK, SINGER, *et al.*, 2008) points out that a major challenge in survey research is to control for sampling bias. Bias can be caused by issues in the representativeness of sample or low response rates. The authors also highlight as another major challenge is to ensure that the questions are well designed.

Observations were also found out in the literature review in studies S04, S07 and S13. (MERRIAM, 2009) explains that observations take place where a given phenomenon naturally occurs. She points out that the skills to be a good observer must be learned, thus, training and mental preparation is important. She highlights the need to define what to observe as well as to write careful and useful field notes. There are different possible relationships between observed and observant varying from complete participant and complete observer. She mentions that combining the role of participant and observer may be a challenge. In qualitative research where the researcher is the primary instrument of data collection, subjectivity and interaction are assumed. However, she mentions that the real question is how the researcher can account for them in interpreting data. S09 used tape-recordings along with a pre-post test method.

The second question was related to the target population of the studies.Table 7 illustrates this information.

Study ID	Target Population
S01	Third grade students of primary education
S02	Kindergarten students
S03	Architecture students
S04	Disabled students
S05	Undergraduate and master students
S06	Elementary students
S07	Children ranging from 8-10 years old
S08	Sixth grade students
S09	Undergraduate students
S10	College students
S11	High school students
S12	Students of architecture and building engineering (the participants were in their 3rd year of the academic course)
S13	Junior high school (grade 2)
S14	Teacher students and in-service teachers

Table 7: Target population of selected studies

Through this table, it is possible to see that most of the applications are designed to K-12 students or at least to students at this age range (S01, S02, S06, S07, S08, S11, S13). Therefore, it is noticeable that only S02

and S13 work with very young learners (up to seven-eight years old).

A large amount of applications are directed to undergraduate students, such as in S03, S05 (in this case, for undergraduate and master students), S09, S10 and S12.

In S04 there was an application designed to disabled people and S14 presented a science to go project that was evaluated with teacher students and in-service teachers.

The following question was about the constructs being evaluated. Table 8 shows this information for each selected study.

Study ID	Constructs
S01	Efficiency (academic achievement), usability and motivation
S02	Three aspects of learning: language ability, creativity level and scientific thinking (assessment)/impact of the use of AR and the robot to children's development and learning and to the role of the teacher/considerations to be made when using such technologies in the children's dramatic activity (either technologically or educationally)/criteria or the selection of the story when using AR and robot assistance
S03	Academic performance improvement, effectiveness, efficiency, satisfaction and motivation
S04	Human factors (wearability and technology acceptance) and pedagogical aspects
S05	Academic performance, usability (effectiveness, efficiency, satisfaction) and motivation
S06	Learning performance (memory, application and comprehension questions concerning formulas and basic concepts of the Chinese library classification scheme, book classification, ability to correctly return books to the bookshelf and correctly permute books according to book call numbers) and learner satisfaction
S07	Degree of knowledge and participant’s satisfaction
S08	Students attitudes, content learning gains and teacher's opinions
S09	Learner's construction behaviors and learning performance
S10	Learning effectiveness, flow experience, the amount of time spent focusing on paintings, behavioral patterns and attitude of using the guide systems
S11	Learning performance in the subject
S12	Usability, degree of difficulty in the use of AR technology on mobile devices for education purposes, assessment of student's academic performance
S13	Learning effect, learning attitude, satisfaction with the software, cognitive validity and cognitive accessibility
S14	The identity of ICT (information and communications technology)/AR-education/changes in learning environment/the innovative approach applied in the process

Table 8: Constructs evaluated by selected papers

Through the table above, it is possible to see that no studies evaluated solely educational aspects. Efficiency, satisfaction, motivation and other factors were also analyzed.

It is noticeable that many applications were under development or had been recently developed, therefore, technology aspects as well as concerns of usability and users' attitudes and satisfaction were present in almost all studies (except for S09, S11 and S14).

It is important to highlight that the construct efficiency had different meanings in the works evaluated. S01 considered efficiency as academic achievement. Other works (S03, S05, S12) considered efficiency in terms of usability as the relation to the expenditure of time and effort for solving the proposed exercise.

Regarding the involvement of the teacher, it was observed that studies involved the teachers in different ways and levels and that some of them did not involve them in the evaluation process (S03, S06, S09 and S10).

Some studies took into account the local curriculum in order to develop content for the tool. For instance, S01 reported that the development of the didactic contents was agreed with the Department of Education of the Comunidad Valenciana (Spain). The AR application has been conceived as a tool for supporting teachers during their explanation, and as an auxiliary resource for the students that can be used for individual learning and to provide a framework for team activities. It was reported that teachers expressed a positive opinion with the experience of using AR contents, and thought that it really had a positive impact on the class. However, it was not very clear in the text how the teacher's opinions were collected.

S2 conducted a Delphi study through three stages. The first one consisted of elements that referred particularly to the possible requirements and the impact of the AR or robot based educational contents and their unique characteristics in relation to children's education. As for the second part, further corrections and improvements were made resulting in 36 final questions. The experts were asked to indicate appropriateness (to the theme of study) of each question in a 5 Likert scale. And in the final survey, the experts reviewed other's opinions and the overall statistical results to reevaluate their own responses. The teachers also had to be trained with the operation control of the robot.

In S04, it was reported no involvement of the teacher. However, it was mentioned that one of the subjects with the arthrogryposis condition used the system with the help of the teacher, because of the bad condition of the shape of their body (stark shoulders and hands).

In turn, S05 also reported no involvement of the teacher. However, it was mentioned that each experimental group has been able to visualize a virtual model created by them or their teachers, in order to evaluate an architectural proposal or a construction detail, on site, as part of their own learning process. In addition, they mentioned that during the sessions, students were able to consult and clarify doubts with the teacher.

In S07, the teachers were involved in a preliminary study carried out in order to determine the subject preferences for an educational computer game for children ranging in age from 8 to 10 years old. The objectives were: (a) to know the opinion of education professionals about the role those new technologies can play in this field, (b) to identify the most appropriate type of game for application in educational games, (c) to identify the most appropriate subjects for

application in educational games, (d) to know the professional's knowledge about AR. The study was designed to reflect the opinions of a large community of professionals from the whole country, and, therefore, the method chosen to gather the data was an anonymous survey. The data collected influenced the design of the game.

S08, on the other hand, investigated teacher's experiences. One of the measures was teacher's judgments of usability and value of technologies related to field trip instruction. They also collected feedback from teacher participants (3, in total) including a group post-interview with the teachers and ecology center program director and individual teacher post-surveys. Prior to the field trip, two of the teachers used learning quests during the class, while the third teacher used them as one of the "stations" during the activities on the day prior to the field trip. During the field trip, the teacher led a discussion about the data collected by students.

The design of the game reported in S11 allows for teacher participation, however, in the experiment they mentioned that one of the researchers acted as the teacher for the session, carrying out a predefined script that was repeated in every session.

S12 mentioned that projects chosen for the experiment were preselected by the academic coordinators and the university studies board of directors. They compared the grades obtained during the academic year (with the AR technology) for all groups and the academic results for all groups of the previous academic year (2010-2011), when traditional methodologies were used. Finally, the authors indicated that this educational research project falls under the Interest Group for Logistics and Teaching in Architecture (GILDA), an inter-university group centered in the architectural framework assigned to the ICE (Institute of Education Sciences at the Polytechnic University of Catalonia – UPC), specialized in the field of teaching technology disciplines.

In S13, researchers interviewed the chemistry teacher before the design and development of the AR tool. The quiz applied was devised by a junior high school chemistry teacher and further examined by a group of chemistry education experts, including two junior high school chemistry teachers and three professors specializing in science education. The designed inquiry-based group learning scenario required students to conduct explorations in groups of three without teacher instruction. They were supposed to use the AR tool and conclude the principles by themselves.

Finally, in S14, the data (N:292) was collected as a sample from 128 in-service teachers and 164 teacher students. The data was collected by Likert-scale questionnaire forming ordinary scale items and factors.

Through the analysis of this data, it is noticeable that most of them involved the teachers or school staff, such as coordinators in development of the tool or project decision as in S01, S02, S07, S12 and S13.

Some studies reported that teachers helped during the execution of experiments as in S05 and S06.

Only S08 investigated teacher’s experiences among the other objectives.

S12 considered the role of the teacher in their game, however, it reported that one of the researchers acted as the teacher during the experiment.

S13 reported to have devised a student-centered experience, therefore, students were expected to explore and reach conclusions through group effort without benefiting from the teacher's direct instruction.

It is also important to point out that some systems were related to education in general, such as S10, which developed a system to be used in art museums. This may justify the fact that there was no teacher involvement in its evaluation.

S14 collected data from both teachers and teacher students and analyzed with the research tool New Educational Models or Paradigm with 27 items.

As for multiple metrics, few studies have used both quantitative and qualitative metrics to evaluate learning gains. Although many studies have used them to evaluate different aspects, such as satisfaction and usability.

We also observed if studies have used different metrics of each of these types (quantitative and qualitative).

Table 9 presents the studies concerning these two aspects.

Study ID	Q1	Q2
S01	Yes	No
S02	Yes	No
S03	No	No
S04	Yes	No
S05	No	No
S06	No	No
S07	Yes	Yes
S08	Yes	Yes
S09	Yes	No
S10	Yes	No
S11	No	No
S12	No	No
S13	Yes	No
S14	No	No

Table 9: Studies concerning the use of multiple metrics. Q1: Did the study use multiple metrics (both quantitative and qualitative)?, Q2: Did the study use multiple metrics for educational purposes?

Another question was related to the applications' domains of knowledge as presented in Table 10. In this table, it is possible to see that the applications were related to the fields of science (S04, S08, S09, S11, S13, S14) and humanities (S01, S02, S03, S05, S06, S07, S08, S10).

Study ID	Application Domain
S01	Knowledge of the natural, social and cultural environment
S02	Dramatic play (children's play)
S03	Architecture
S04	Science (physics)
S05	Architecture (construction and maintenance learning process)
S06	Library instruction
S07	Multiculturalism, tolerance and solidarity (focus on transmitting knowledge about three of the world's poorest continents (Africa, Asia and Central and South America)
S08	Ecosystem learning
S09	Science: Elastic collision and momentum
S10	Art appreciation instruction
S11	Collaborative game (electrostatics)
S12	Architecture
S13	Chemistry - the composition of substances
S14	Science

Table 10: Applications domains

Through these results, it is noticeable that there is a lack of studies regarding early literacy development using AR systems.

The research questions of the studies were classified as explicit or inferred as they were explicitly described in the text or not. Figure 18 shows that half of the papers presented explicit questions and the other half, inferred ones.

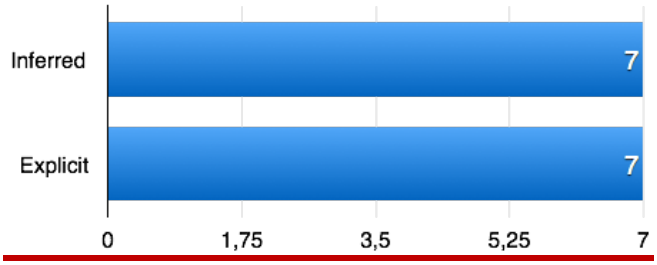


Figure 18: Question types: explicit versus inferred questions

The questions were also classified according to their types as proposed by (EASTERBROOK, SINGER, *et al.*, 2008). The authors explain that there are different types of questions as will be described below:

- Exploratory questions: according to authors, this type of question is asked in the early stages of a research program when researchers are attempting to understand the phenomena. Examples of this type are: existence questions, description and classification, descriptive comparative.
- Base-rate questions: this type is frequently asked after having a clearer understanding of the phenomena. These questions might be: frequency and distribution questions and descriptive-process.
- Relationship questions: these are meant to understand the relationship between two different phenomena.
- Causality questions – these questions are an attempt to explain why a relationship holds and identify its cause and effect. These questions can be: causality questions, causality-comparative questions and causality-comparative-interaction questions.

The aforementioned questions are named knowledge questions. However, the authors explain that software engineers need to ask also a different question concerned with better ways to do software engineering, the design questions.

Figure 19 presents the types of research questions found in the selected papers.

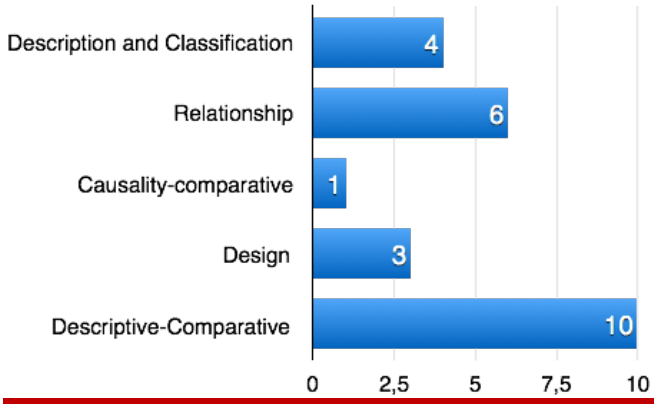


Figure 19: Types of research questions

Some papers presented more than one question. Through the chart above, it is possible to see that most of the questions (10, in total) were descriptive-comparative. Those papers intended to describe the effect of a given AR technology comparing it with different resources.

The second most common type was the relationship question (6, in total). In this case, papers wanted to investigate the relationship of AR usage in different aspects, such as academic achievement or motivation.

Description and classification questions were also found (4, in total). Those papers aimed to describe and classify the experiences with AR systems applied in educational settings. Design questions were present in 3 papers. Those questions intended to investigate issues such as the

feasibility of using AR in a specific context or assess particular provisions needed both educationally and technically.

Only one paper asked a causality-comparative question. This paper intended to investigate if the learners who learned with the AR tool presented better learning achievement than the ones who learned with traditional 2D technology.

Another classificatory question of this systematic review regarded the AR technology used in the studies. AR technologies were classified according to the tracking, display and interaction techniques employed. The results from this question are presented in Figure 20.

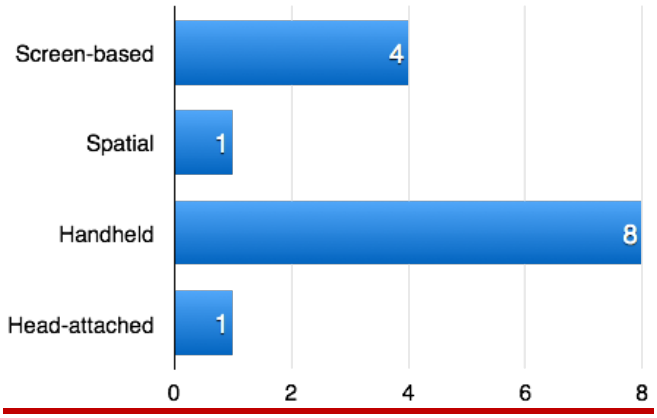


Figure 20: Papers according to the display techniques used

By analyzing this figure, it is possible to see that regarding display techniques, handheld ones were the most common choice of the papers analyzed (8 papers, in total). The popularization and technical advancements in smartphones make handheld displays a good option for AR applications. These types of devices are minimally intrusive and highly mobile as pointed out by (ZHOU, DUH and BILLINGHURST, 2008).

We noticed this is a positive characteristic since it allowed for experiences in different settings such as field trips (S08), museums (S10) and many different places, which can foster independence on learners. For instance, S07 also emphasizes the versatility of the learning activity (in this case, an iPhone game) that can be played at any place and time without requiring supervision.

On the other hand, S11 reported that their AR system was considered more costly than the traditional one. This may stand from the fact that in many schools is more common to have desktop-computers rather than mobile devices.

Another limitation was pointed out by S12. This study mentioned that the AR system showed to be good to visualize simple models but not complex ones, either in volume or structure. They obtained low reviews for the obtained perceptions of the visualization of complex models, not only architectural projects.

Screen-based displays were the second most used type in the selected papers (4 papers). This type of display is known for its cost-efficiency since it requires off-the-shelf hardware and standard PC equipment. It is also largely present in schools nowadays. Those displays were usually well evaluated by users. For instance, S06 highlighted that the system was helpful in promoting learner motivation and that they were satisfied with the tool. S13, in turn, reported that students generally had positive attitudes toward the software.

Head-attached displays were used in one of the papers. S04 used head-mounted displays with disabled students who had different levels of mobility difficulties. Results showed that disabled students had almost the same results as the able-bodied ones. However, it was pointed out that the system

was considered fatiguing and also fashion constraints were present. Although these displays provide a better field of view and, as exemplified by S04, may foster inclusion, fashion constraints are commonly reported in the literature.

Spatial displays were also present in one of the works and as opposed to head attached ones they seek to eliminate the need for equipment attached to user's bodies. These displays project virtual content directly on site. Thus, it allows multiple users at a time and collaboration. This characteristic of AR (to allow virtual objects appear registered in the right place in the real world) was pointed out by S02 as particularly opportunistic and suitable for supporting children's play. Besides that, it was pointed out that as a novelty, it may draw student's attention, replace physical masks and costumes with digital augmentation and using props as means for all-purpose interaction. Therefore, that was the choice of authors to their system which showed positive outcomes.

As regards to tracking, most of the papers presented vision-based tracking (13 papers, in total). One paper presented hybrid-tracking as presented in Figure 21.

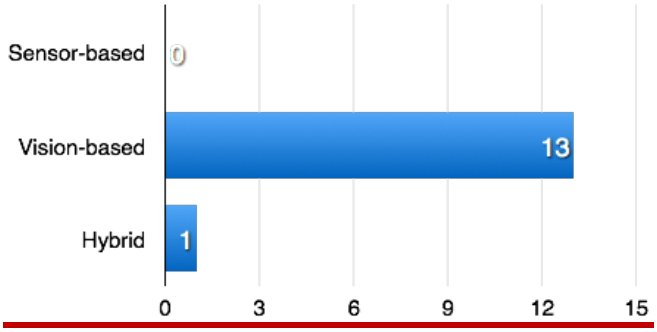


Figure 21: Papers according to the tracking techniques used

Vision-based tracking can be divided in two different categories, marker-based and markerless. According to this classification, 11 of the selected papers presented marker-based tools while 2 papers presented markerless tracking as illustrated in Figure 22.

Thus, marker-based was the most common type found in the papers. Marker-based is a very popular type of marker since there are many marker-based kits available. Most of the papers presented positive outcomes regarding these tools. However, some papers highlighted as a negative factor the need to alter the environment with fiducial markers and the need to previously locate them in the scene (S03, S11). Additionally, S11 pointed out that the AR tool was more complex to organize since it was necessary to arrange the desks to provide space for movement by students, locate the markers and adjust lighting conditions whenever necessary. Another issue reported in some papers was instability of the scene (S02, S03, S13) due to different factors such as lighting conditions, occlusion of markers, changing distances, angles between cameras, among others.

On the other hand, markerless systems do not require he use of markers in the environment. In this case, the environment itself act as marker. This type of tracking was chosen in three papers. These works generally presented positive results. For instance, S08 pointed out that AR was most effective as a mode of engagement and as a way of structuring and enhancing the probeware-based activities of a field trip.

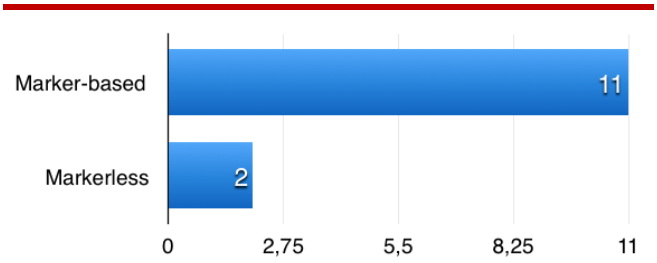


Figure 22 Papers according to the vision–based tracking used

Finally, as regards to interaction techniques, most papers presented a more traditional type of interaction using buttons, touch or simply providing visualization of the augmented content (7 papers). Their use was generally positive. For instance, S13 reported significant supplemental learning effect of the AR tool and other reported that their systems generated equivalent learning to the alternative option they were compared with (S06, S07). S07 reported that boys outperformed girls when using the system while in S08 no gender difference was observed. S10, on the other hand, reported that the AR guide effectively enhanced visitor's learning effectiveness, however, they mentioned that the AR guide was found to promote less peer interaction.

The second most common choice was collaborative interaction (4 papers). These types of applications allow both face-to-face and remote collaboration by integrating multiple user's devices in different contexts and, therefore, enhancing telepresence.

The outcomes were generally positive, supporting student's independence (S08) and generating an active student (S12). Positive results were also observed regarding student's satisfaction, motivation and engagement (S12). For instance, S08 pointed out that the AR tool tested enabled collaborative communication and problem solving among students.

It is interesting to highlight that S09 aimed to investigate the impact of a mobile AR system to support learner's collaborative knowledge construction processes and enhance this learning achievements regarding the topic of elastic collision. This work presented positive outcomes since AR learners presented significant better learning and the tool supported the student's knowledge construction processes. However, it was noticeable that students would sometimes seek for further clarification of the theoretical principles or concepts. According to the authors, this result might indicate that while students constructed relations between single theoretical concepts or distinguished concepts from each other, they would confront obstacles regarding the topic of elastic collisions and need support from their partners. They also acknowledged that it might be possible that they acquired inaccurate concepts from the other dyad learners. Therefore, they recommended as future improvements to add adequate representational guidance to support learner's cognitive behaviors and specify the relationships between the concepts.

Tangible interfaces were chosen in two papers. These interfaces are promising since they take advantage of the familiarity of everyday objects to ease the interaction. The use of those types of displays provided positive results. For instance, S01 pointed out that students and teachers showed positive attitudes towards the tool. Additionally, they mentioned that student's found the AR tool ease and natural to use.

In turn, S14 mentioned as positive that with AR it was possible to combine real objects with virtual ones and to place suitable information into real surroundings. In

addition, they mentioned that the use of AR allowed students to interact physically and intellectually with instructional learning scenarios material through "hands on" experimentation and "minds on" reflection as a result the pedagogical experts and teachers attending the process underlined as the main element moving from teacher controlled learning to student oriented learning with context-related knowledge.

Hybrid interfaces, on the other hand, were used in one of the works. As previously mentioned, hybrid interfaces differ from others since it combines different but complimentary interfaces and displays as well as the possibility to interact through varied interaction devices. S02 combined AR to robotics in order to support children's play. This combination was evaluated as positive by authors who highlighted that the robot not only assisted the teacher but also implicitly interacted with the audience in promoting their immersion into the play as well.

The results regarding the interaction techniques used in the papers can be seen in Figure 23.

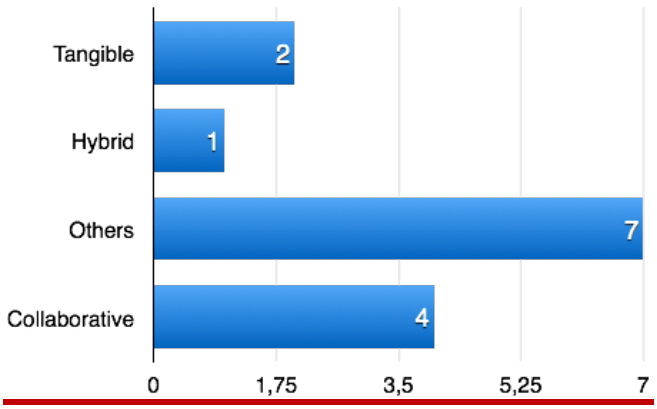


Figure 23: Papers according to the interaction techniques used

Concerning the AR definition used in the papers, most of the papers presented the definitions given by (AZUMA, 1997) and (MILGRAM and KISHINO, 1994) as in S01, S06, S07, S14. S12 mentioned only the latter definition.

Some of them mentioned (MILGRAM, TAKEMURA, *et al.*, 1994) such as: S01, S09, S11.

The other papers did not mention AR definitions or mentioned varied references in order to exemplify AR's applications or highlight some of its characteristics (S02, S03, S04, S05, S08, S10, S13).

Thus, it is noticeable that there is a consensus regarding the AR definitions used in the papers.

Another aspect observed was if the selected papers based their work in any educational theory. Table 11 presents the results for this question.

Study ID	Educational Theory
S01	Does not mention it
S02	Robot assisted learning
S03	Does not mention it
S04	Does not mention it
S05	Mobile learning
S06	Situated learning theory
S07	Experiential learning theory/theory of multiple intelligences by Gardner
S08	Situated learning theory
S09	Technology-enhanced learning/collaborative learning
S10	The system was based on mobile guide design principle of the HCCI (Human Computer Context interaction) framework and art appreciation theory.
S11	Computer supported collaborative learning
S12	Mobile learning
S13	Inquiry-based learning

S14	Experience based-learning /IBSE (inquiry based science education) model and 5E (inquiry approach)

Table 11: Papers according to the educational theory used

Through this table, it is possible to see that most of the papers mentioned educational theories (10 papers, in total). Among the educational theories mentioned are: robot-assisted learning, situated learning theory, experience learning theory, theory of multiple intelligences, technology-enhanced learning, collaborative learning, art appreciation theory, computer supported collaborative learning, mobile learning, inquiry-based learning, experience based-learning, IBSE model.

Something that most of these theories have in common is a student-centered approach. Some of these theories also discuss the use of technology in the learning process.

Finally, it was investigated if the AR applications were combined with different technologies and what kinds of technology they were combined with.

10 studies did not combine AR with other types of technologies. The other studies combined AR with different technologies and evaluated them. S02, for instance, combined AR technology with a remote controlled-project-camera (pro-cam) robot. In S04, the AR connect project uses a mobile AR technology based system and an associated computer-mediated learning platform. S07 combined AR mini-games with non-AR mini games. In turn, S08 combined AR to handheld environmental probes.

It is important to highlight that these combinations of technologies provided positive outcomes. For instance, in S02 authors reported that the robot associated with AR technologies relieved the burden of coordinating the play and running the AR system. In addition, they highlighted that it served as another attention drawing entity, resulting in synergically enhanced educational effects.

S04 reported that disabled students had almost the same results as the able-bodied ones in the tests. However, the observation suggested that the system was fatiguing to wear regarding some aspects.

In S07, results indicated that children achieved similar knowledge improvements using an autonomous game (iPhone game) as when using the traditional alternative. However, authors did not check the contribution of each feature for the outcomes presented.

S08 reports that the combination of AR and the probeware helped to situate the measurements in a meaningful context as students were able to carry the data they had collected back into the classroom. Their results also showed the activities which integrated probeware resulted in significant learning gains related to student understanding of water quality variables.

4.4.4. Relation And Effect Questions

The questions in this category were: (10) what are the implications of the findings in research and practice? and (11) what is the impact of the tool analyzed (positive or negative)?

Through the analysis of question 10, it was possible to observe that AR has implications in many different areas and there are also areas for further exploration.

For instance, AR showed to be useful to enable a more direct and active involvement of students (S01).

The combination of AR with different technology also showed to be a relevant aspect (S02, S07, S08). In S02, AR was combined with robotics. In this work, the robot relieved the burden put on teachers of coordinating the play and running the AR system. In addition, it served as another attention drawing entity, resulting in the synergistically enhanced educational effects.

AR technology combination with mobile phones showed to be very promising as pointed out in S05. This study showed that AR in combination with mobile phones offered many possibilities to evaluate, on site, architectural projects, urban design, construction processes, and historical heritage studies. It facilitated social dissemination by showing the real scale and position of historical buildings in real time.

Many works have shown impact on student's motivation (S03, S08, S10, S12)

In S03, it was also pointed out that the AR technology had some advantages over commercial applications, such as the possibility to display several 3D models without changing the marker and ability to move objects in the scene as well as the viability of outdoor studies.

AR showed to have an impact in different fields of instructional technology, educational systems design, science education, among others.

Using AR technology also imposes some challenges, for instance, S04 noticed that their AR experimental scenario should be much more carefully designed and take under consideration the background of students (e.g.: living conditions, previous knowledge).

AR also had impact in instructional activities, S06 pointed out the possibility of the AR technology to replace conventional library instruction.

There are also a number of studies that discussed the possible impact of AR technology depending on a number of factors, such as gender (S06, S07, S11), student's personal learning styles (S06), personal gaming skills (S06) and student's interest in technology (S12).

AR also enabled to explore learning outside classroom doors (S07). Additionally, it enabled to teach different contents, such as multiculturalism, solidarity and tolerance.

S07 considered formal educational learning theories in their evaluation and compared their system with traditional methods. Therefore, this study emphasizes the importance of using pedagogical foundation in order to design educational computer games and the importance of using control groups and considering national curricula in the development of those games. This study also claimed to be the first one to take into account the preference of professionals.

AR can also allow children to explore what they are learning from a variety of perspectives. Studies emphasized the possibilities of combining AR with other technology (S02, S04, S07, S08).

S07 also emphasized the importance of the use of multiple interaction forms in the game (e.g.: touchscreen interaction and accelerometer). Some possible future works of this research are related to the need to allow more involvement of the teachers, e.g.:

letting them establish the game difficulty taking the level of students into account.

AR has shown to enable the creation of student-centered learning experiences. In addition, it provided opportunities for peer teaching, collaboration and one-on-one teacher guidance (S08). AR also enabled learning opportunities that could not be possible without the technology. It also allowed the use of pedagogical approaches that may otherwise be difficult in an outdoor learning environment besides it supported independence for students.

Another issue that has been investigated is learner's behavior patterns and their knowledge construction processes (S09, S10). Design and usability aspects also have great impact on AR systems and were investigated in many studies (S01, S04, S05, S08, S12). S10 compared visitor's behavior patterns in a museum considering the type of guide used, AR-guided, audio-guide and no guide. Results showed that visitor's patterns were dependent upon the guide mode used. Most visitors believed that AR guide made it easier to digest information than the audio-guide due to extra visual commentary that was provided. Results of behavior patterns indicated that the AR-guided mode deepened the interaction between the viewing of the artwork and its explanation. The analysis of behavior patterns also revealed that during the activity of painting appreciation with the system, the visitors did not focus on the device excessively, and, thus, did not largely ignored the paintings. Nevertheless, it was found out that the AR-guide promoted less interaction with peers.

The system described in S10 is not only limited to painting displays, but it may be applied to all kinds of exhibitions (e.g: museums and theme parks).

Another aspect of great relevance is cost. For instance, in S11 pointed out that their AR solution was more costly than the traditional one.

In turn, S12 investigated good practices for technology acceptance as well as how to implement new teaching methods with mobile technologies. Authors highlighted the importance of students creating information as well as the need to work collaboratively. In this study, the importance of using familiar technologies, such as mobile devices to visualize architectural models was emphasized.

S13 discusses the implication of AR in students regarding their academic achievements (if they were either low-achieving or high-achieving students). It was shown that the AR incorporation had a larger influence on low-achieving students. This research showed that like any other learning tool the gains AR may produce are based on student's belief that learning the discipline is important. Future works suggested involve the investigation of AR as a remedial tool.

Another implication for AR is that it may serve as a bridge to formal and informal learning as presented in S14. In this work, the use of AR allowed hands on experimentation, minds on reflection and the move from a teacher-controlled to student oriented learning with context-related knowledge. This work raised interesting implications such as that AR enables the combination of education and thinking skills. Additionally, they highlighted that teachers were not impressed by technology itself but for the connection between learning environments.

The second question investigates the results of the researches, if they were either positive or negative.

Most of the results were positive. Many applications presented positive outcomes regarding learning and academic performance (S01, S02, S03, S05, S09, S10), student's motivation and engagement (S03, S05, S06, S07, S08, S10, S12).

Many studies also presented positive outcomes regarding usability aspects (S01, S05). Positive outcomes were also found out regarding aspects such as efficiency, effectiveness and satisfaction (S12, S13).

In S04, disabled students achieved almost the same results as the able-bodied students after using an AR tool. Authors, however, argued that results were not conclusive since the sample size used was limited to provide statistical analysis. As it comes to usability aspects, they found out that the system was fatiguing to wear which led to redesign of the tool.

S06 presented a somewhat neutral outcome regarding learning performance. Authors argued that the proposed ARLIS and librarian instruction for library instruction generated equivalent learning performance. The authors, however, presented positive results regarding motivation and willingness to learn. This work also reported that the system is more helpful in promoting the learning performance of learners with the field-dependent cognitive style than the conventional librarian instruction, particularly for learning content associated with application and comprehension. The work confirmed that the learning performance was not affected by personal gaming skills.

S14, on the other hand, concluded that AR is a promising technology for classrooms and that the possibility of AR to make convergence of education is challenging. The AR system evaluated in this study enabled student's active participation as a result the pedagogical experts and teacher attending the process underlined as the main element moving from teacher-controlled learning to student oriented with context-related knowledge. They also concluded that usability, availability and the prices of this AR-technology are making it soon available for everyday education routines.

AR also showed to have an impact on user's behavior patterns. In S10, researchers compared three different situations (an AR based guide, an audio-guide and no guide) in guiding visitors through a museum. Results showed that visitor's behavior patterns were dependent upon the guided mode used.

The combination of AR with different technology was also promising, for instance, S02 presented a system combining AR with robotics in order to manage children's play. In this study, the robot relieved the burden put on teachers of coordinating the play and running the AR system. In addition, it served as another attention drawing entity, resulting in synergistically enhanced educational effects. S08 combined AR with handheld environmental probes. This work argued that AR was most effective as a mode of engagement and as a way of structuring and enhancing the probeware-based activities of the field trip.

It was also noticeable that AR may have an impact in a wide range of situations with varied age levels and different contents ranging from physics and architecture to multiculturalism and solidarity.

AR was also said to support independence (S08), active student (S12) as

well as to promote “hands-on” experimentation and “minds-on” reflection (S14).

It was noticed that AR also offered some constraints. For instance, S03 reported some limitations regarding stability, which were: strong dependence of ambient light conditions, camera distance and marker size relationship, and the fact that the physical environment was altered by fiducially markers. Stability problems were reported in other works as well. S02 pointed out that from time to time the AR system failed due to different factors, such as: changing in lighting conditions, occlusions of the marker, changing distances or angles between cameras, among others. Nevertheless, this did not affect children who after somewhat understanding the operation, showed strong interest and persistence to work with the system.

S04 discussed the question of fashion constraints. In this work, some participants were self-conscious or embarrassed to wear the system. Some were stressed out or reported fatigue.

In S07, although general results were positive and scores were quite high in the two situations tested (an iPhone game versus a traditional one), people in charge of the study noticed that players understood the concepts better when using the traditional method than when using the iPhone version.

In turn, S08 reported some teacher’s concerns regarding the tool tested, which were: (a) managing the tension between positive aspects of student engagement and student’s desire, negative in its effects on learning, which is a common problem in any field trip and (b) concerns about their ability to manage the technology and devices when orchestrating the field trip on their own. As pointed out in the research, during the experiment teachers had the help of researchers during the entire process.

Another important aspect is the interaction with peers. S10 reported that their AR-guided mode for painting appreciation promoted less peer interaction.

Costs were also an issue to be taken into account, for instance in S11 the AR proposed solution was more costly than the traditional one which discouraged its use. In this work, it was also pointed out complexity issues since the AR platform required a more extensive setup (i.e: arranging the desks to provide space for movement by students and locating markers and adjusting lighting conditions when needed).

In turn, S12 reported doubts about the adequacy of AR for complex project realization, which was compounded by the low reviews, obtained of the visualization of complex models.

4.5. Final Remarks

Through this research, we could identify AR’s potential to be applied in classrooms. Developments in AR technology have enabled researchers to develop more tools in the field of education and to evaluate them. Therefore, it was evident that there is also a growing interest in evaluating its impact in the learning process.

These findings were used as a solid literature foundation for our evaluation. Thus, it is important to position our work in this scenario.

Through the investigation of the methodology design applied in the evaluation, it is possible to observe that most

of the studies combined different methodologies in order to evaluate their tools, although, only few of them combined them to evaluate educational gains.

Regarding the use of multiple metrics, we found out that only few works applied multiple metrics when it comes to evaluate learning effects. Most of them do use multiple metrics but in order to evaluate different aspects than learning, such as usability, efficiency, motivation and user's satisfaction. (MERRIAM, 2009) explains that all research designs can be discussed in terms of their relative strengths and limitations. She points out that their merits are related to select the most appropriate ones to address a given research problem. Thus, (COHEN and MANION, 1989) explains that the advantages of using multimethod approach in social research are manifold. The authors highlights two of them: (a) while single observation in fields such as physics and chemistry usually yield sufficient and unambiguous information, it provides a limited view of the complexity of human behavior and interactions and (b) exclusive reliance on one method, may bias or distort the researcher's picture of a particular reality he/she is investigating.

It was also observed that few studies involved the teacher as an instructional designer. Nevertheless, some studies took into account the local curriculum in order to develop content for the tool or took into account expert's opinion in the development process of the application and also during the development of evaluation material. (FITZPATRICK, 2004) highlights the need to involve teachers in the process of adopting new technological tools so activities are integrated into their lesson plans and, thus, are meaningful to the students.

In our evaluation of an AR educational system, the ARBlocks, we also follow the tendency to combine different methodologies in order to have a more accurate feedback of its use. Therefore, we will employ a quasi-experimental design, followed by interviews, a research-diary (this instrument was not used in the pilot study as will be appropriately explained in chapter 5), a post-test and observations. Moreover, multiple quantitative and qualitative metrics will be employed during the evaluation process in order to evaluate educational impact. This work will also take into account during the evaluation the teacher's point of view and evaluations in order to have a better overview of the technology insertion process.

Since teachers are the experts in their students needs, the evaluation carried out in this research will involve them in the development of content for the tool as well as in the evaluation process. Hence, only a post-test will be applied since the contents worked with the tool will be shaped along the experiment. In order to know the previous level of students the middle term evaluation applied by the teacher will be used.

As it concerns to the development of the ARBlocks, researchers based this process on solid research regarding its design and the context for which the tool was aimed. It was adopted the user interface design in order to bring teachers closer to the process of designing the tool as well as collecting feedback about it (ROBERTO, 2012).

In the review, it was also noticed that there are solutions being developed to different age groups and knowledge domains. However, it was noticed a lack of

evaluation of AR systems aimed at very young learners. The present work intends to provide some additional experience on that topic since the tool evaluated, the ARBlocks, as will be later detailed, was developed aiming at young learners.

Our study will encompass mostly educational aspects although satisfaction and motivational issues will also be taken into account during the experiment. Some dimensions of the process, such as teacher's perceptions and how they adapted the tool in their classroom will be considered as well. When it comes to the application domain, it was found no work in the same domains as ours which is early literacy. It is important to highlight that the field of early literacy is a very challenging one, especially when it comes to learning a foreign language. Additionally, it is important to highlight that designing and testing technology tools with young children has its own characteristics and challenges as pointed out by (DRUIN, 2002).

Thus, our work intends to provide some insight on the process of evaluating AR systems in the scenario of early literacy development as well as reflect about the challenges involved in this process.

Regarding the types of questions asked in the selected studies it was noticed that most of the questions were descriptive-comparative. Many papers intended to describe the effect of a given AR technology comparing it with different resources.

The second most common type was the relationship question. Those papers aimed to investigate the relationship of AR usage in different aspects, such as academic achievement or motivation.

Another type of question found out was the description and classification question which intended to describe and classify the experiences with AR systems applied in educational settings. Design questions were present in three papers that aimed to investigate issues such as the feasibility of using AR in a specific context or assess particular provisions needed both educationally and technically.

Causality-comparative questions were found out in only one paper that intended to investigate if the learners who learned with the AR tool presented better learning achievement than the ones who learned with traditional 2D technology.

As for our evaluation work, we investigate description and classification question as its objective is to investigate the effects of AR in the process of evaluating educational technology holistically involving both teachers and learners in the process. Relationship questions were also investigated since it is also in the scope of our work to evaluate the use of the ARBlocks in the language learning field concerning the following aspects: linguistic concepts and competences.

The questions chosen in our case aim to explore how the AR tool impact in the learning situation in the scenario chosen as we attempt to understand what are the implications of using the ARBlocks in the language teaching environment. We used a control group in order to compare student's academic evolution with and without the use of the tool. Although we recognize that would be interesting to compare an AR tool with a different resource, it was not possible in our case to isolate the external factors and the methodology used in the school in other

groups. Therefore, we compared the AR tool with the other group receiving the conventional lessons applied in the school. Although it is important to realize that the control groups were provide with practicing the same content using other resources.

Usability issues were not part of our scope since we understand that these issues regarding the tool evaluated were already covered in previous works (ROBERTO, FREITAS, *et al.*, 2013).

As part of our systematic literature review, the papers were classified according to the tracking, display and interaction techniques employed in their systems.

As concerns the technology choice of the tools analyzed, it was noticeable that this choice of technology varied deeply depending on the learning objectives of the tool. Nevertheless, this choice had an impact in the possibilities as well as the limitations of use of a given AR application.

As regards to this aspects, the system evaluated in this work, the ARBlocks, is classified as using spatial displays since it enables the projection of virtual content on different surfaces such as a table or the floor. It is a marker-based system and enables tangible interaction with the content. This last characteristic is especially important for young learners who benefit greatly from the interaction and manipulation of real objects.

Another aspect observed during this literature review was the AR definition used in the papers. Most of them presented the definitions given by (AZUMA, 1997) and (MILGRAM and KISHINO, 1994) which were the same definitions used in this work.

The other papers did not mention AR definitions or mentioned varied references in order to exemplify AR's applications or highlight some of its characteristics.

Another aspect observed was if the selected papers based their work in any educational theory. It was observed that most of the papers mentioned educational theories (9, in total).

However, it was noticeable that some papers (4, in total) did not mention educational theories. We believe that educational theories may help to unravel contributions of AR tools as well as its limitations. In addition, it may help to understand how AR unique features may impact in the learning setting.

As for our work, the tool evaluated is based on the educational works of Piaget (see section 3.1.4) as this author works with young children development (ROBERTO, 2012). Our evaluation was based on solid research regarding language learning and children's development as previously exposed in chapter 3.

It was also investigated if the AR applications were combined with different technologies and what kinds of technology they were combined with.

For this aspect, the majority of studies did not combine AR with other types of technologies. The other studies combined AR with different technologies and evaluated them.

Our work follows the trend found in the majority of the studies and evaluates solely the AR system, the ARBlocks.

The last two questions of our systematic review concerned the implication of the results for practice in the field and the quality of the outcomes (if they were either positive or negative).

Regarding the former question, it was observed that the results obtained were highly related to the types of questions asked.

Thus, many studies discussed the impact of AR tools in academic performance and learning opportunities enabled by AR systems. Some works have also discussed mobile learning opportunities created with AR.

In addition, there were works discussing improvements and redesign of the tools being evaluated. Others discussed the AR influence on low-achieving students.

The latter question investigated the results of the researches, if they were either positive or negative.

Most of the results were positive. Many applications presented positive outcomes regarding learning and academic performance, motivation, engagement, usability, efficiency and effectiveness.

There were papers that found usability issues that led to the redesign of the tool.

There was also a neutral outcome regarding learning performance which means that the proposed AR system generated equivalent learning performance when compared to a traditional one. This work, however, presented positive results regarding motivation and willingness to learn.

As for the evaluation conducted in this study, it is expected that the introduction of AR technology will benefit students learning, therefore, providing positive results for the subjects involved in the process.

The actual results will be appropriately discussed later on in this work in chapter 7.

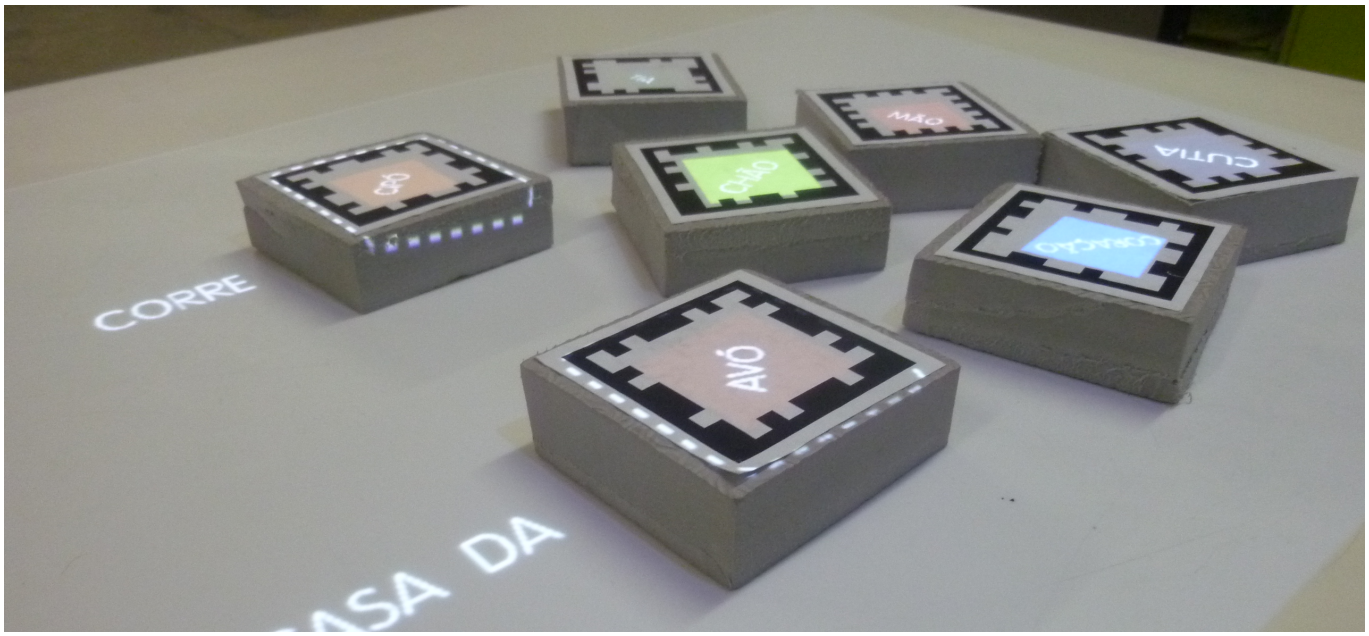
To sum up, during this systematic review, it was noticed that AR has unique affordances that can impact the learning experience. As technology matures, researchers are increasingly concerned with how to incorporate real classroom issues into their investigation.

The results of this review helped to shape and to position our own evaluative study (that will be appropriately described in the following chapters in the literature scenario of research.

As for limitations of our review, due to the limited number of databases evaluated, authors are aware that results may not fully represent the development of research in the field.

5. METHODOLOGY

Figure 24: Nursery rhyme activity. In this activity, students were required to complete the nursery rhyme with the missing words displayed in the blocks



(EASTERBROOK, SINGER, *et al.*, 2008) point out the usefulness of mixed methods in the research design. These authors highlight the importance of employing both quantitative and qualitative metrics as a way of compensating the weakness of each method.

In order to effectively evaluate new educational technology it is important to effectively integrate them in the schools. (DEXTER, 2002) points out two premises for effective integration and implementation of technology for K-12 classrooms, that are: (i) the teacher must act as an instructional designer, planning the use of technology to support learning and (ii) schools must support teachers in this role.

Although primarily thought for effectively integrate technology in a long-term period, we believe its important for researchers and developers to have a sense on how teachers will integrate new technologies into their lessons since this will shape student's learning opportunities.

In addition, as regards to evaluation, (CROMPTON, 1996) explains that there are two possible stages for evaluation of information technology (IT):

- The evaluation of the IT intervention in isolation, and;
- The evaluation of the IT intervention within the course itself.

The author explains that the evaluation of a piece of technology in isolation will tend to focus on various aspects of the technology itself, such as screen design and text layout.

On the other hand, the evaluation of a courseware within the course itself will allow for examination of other factors that will lend to successful integration of the product

within the course. Some of these aspects are:

- Educational setting;
- Aims and objectives of the course;
- Teaching approach;
- Learning strategies;
- Assessment methods;
- Implementation strategy.

This work is oriented to provide additional data on the effects of AR in education and on the process of evaluating educational technology holistically, involving both the teachers and the students in the process. In addition, provide some understanding on how to use the aforementioned principles to evaluate the ARBlocks, a projective AR tool especially designed to teach young children.

This work aims to evaluate the use of the ARBlocks in the language teaching field concerning the following aspects: linguistic concepts and competencies.

It also intends to discuss possibilities of use for AR tools in the classroom environment as well as reflect about the potential and difficulties involved in the introduction of this new piece of technology in the school environment and how teachers can adapt this new tool into their teaching routines. Multiple metrics both quantitative and qualitative were used.

Finally, we intend to provide some guidelines in order to assist researchers when conducting similar evaluations.

Firstly, a pilot study was designed in the language learning scenario. Finally, after the lessons learned from the pilot study, a quasi-experiment was conducted in an English learning as a foreign language scenario with two different groups. These studies will be properly described below.

5.1. ARBLOCKS

The ARBlocks is an AR tool developed in order to scaffold education. It combines the principles of projective AR with tangible interaction as presented in Figure 25 (ROBERTO, 2012).

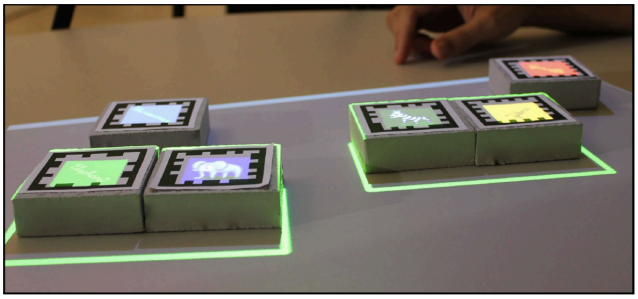


Figure 25: The ARBlocks consists of blocks made of ABS plastic that can have different content projected on their surface. Visual feedback can be offered when the activity is correctly performed

According to the authors, the choice for projective AR enabled the development of a low cost interactive and dynamic tool. In addition to that, the use of tangibles was carefully chosen in order to foster students' learning.

Studies have shown that young children learn better by doing and manipulating realia. For instance, blocks are already part of children's education. Wooden blocks are commonly used in kindergarten classes to learn a plethora of subjects, including language and mathematics. The ARBlocks is an alternative to those blocks allowing for a variety of contents to be projected in its faces. The tool also enables visual or auditory feedback. This subsection will discuss the development of such a tool as exposed by its creators (ROBERTO, 2012). Design and technical aspects will be discussed as well as the reasons why this tool was chosen to be evaluated in our study.

5.1.1. Design Considerations

In (ROBERTO, 2012), creators of ARBlocks explain design issues taken into account in order to develop the application. The design process of the platform was the cartesian method (ROBERTO, FREITAS, *et al.*, 2013). After finding out a problem, which was to develop a low-cost AR application to enhance student's learning, authors tried to solve three main design issues, that are shape, typography and material, in order to build an effective tool both in terms of usability and ergonomics.

First of all, it was performed a benchmark analysis. Authors analyzed regular toy blocks, their characteristics, such as shape, size, materials and how they were used for educational purposes. Typography of contents displayed in those blocks was also analyzed. Researchers discovered that it was important to offer variety. Both print writing and handwriting are important during children's literacy development. Fonts were chosen based on their size, weight and style.

Ergonomic factors were also taken into account. Sound research was considered when designing the blocks' size. Blocks dimensions were specially designed to fit children's hand. Paper drafts were created in order to compare relevant characteristics, namely size and shape. Blocks height was carefully chosen so children are able to manipulate them without occluding the marker. Final solution was designing the blocks with 6x6 centimeters for the top face and 2 centimeters high. The material chosen was the ABS plastic which is harmless and commonly used in children's toys. This material is also cheap which reduces costs of the blocks production. The design process can be seen in Figure 26.

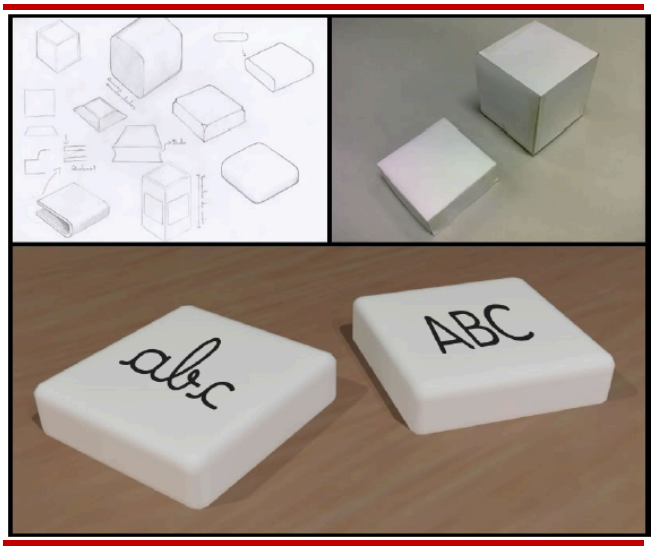


Figure 26: ARBlocks' design process: initial sketches, preliminary paper mockup and final solution rendered with a 3D modelling software

5.1.2. System's Development

After all the design issues were properly handled, the next stage was the system's architecture and computational development. Final users' needs were taken into account at this stage.

The authors explain that the most important premise they must follow is that the interactions needed to be as simple as possible, preferably using the blocks since teachers are not always familiar with computer technology. In addition, from children's point of view, exploring the physical blocks enables them to take advantage of using manipulatives in their learning process.

Another important aspect is the interaction among blocks in order to provide feedback for the students. The system is able to determine blocks' positions relative to one another as well as track information from the blocks, such as where they are in the projection area or their rotation angle.

The ARBlocks' interaction was crucial to select its visualization method, projective AR. This was the most suitable choice for the project since HMDs would not be comfortable to young children and monitors did not allow natural interaction (e.g.: manipulate an object and look to another direction, the monitor, to see its effect).

One positive aspect of the ARBlocks is that different activities can be done with the same set of blocks. The ARBlocks was designed as a platform where educators could choose the activities they want to use as well as propose new ones. However, the tool still lacks an authoring tool to enable teachers to fully explore this characteristic.

In order to make easier for programmers to develop new applications for the ARBlocks, the system was designed as a framework in which the developers would have access to all its basic functionalities (e.g.: tracking or camera-projector alignment). Thereby, the programmer just needs to develop applications that invoke these functions and incorporate them to the ARBlocks as a plug-in.

All these requirements were used to define the relationship between modules proposed in the ARBlocks architecture as can be seen in Figure 27.

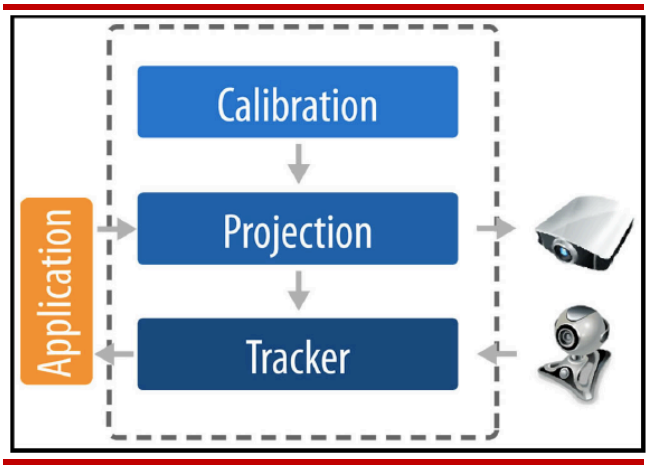


Figure 27: ARBlocks' architecture

The first module to be executed is the *Calibration* which is responsible for making the alignment of the camera–projector system and to deliver this information to the *Projection* module, the last one to be executed. In the main loop, the camera captures an image that will be sent to the *tracker* module. This module finds all blocks in the scene as well as their positions and orientations.

The authors clarify that the core structure of ARBlocks is the combination of these three modules. The last one, the application, indicates what happens when blocks are moved, rotated or placed next to others. This module is also responsible for informing what should be displayed on each block as well as managing the activity. The processing result is sent to the projection module, which is combined with the alignment data in order to project the information correctly on the blocks.

5.1.3. State of the Art of the ARBlocks

One of the reasons this tool, although still a prototype, was chosen is that it is ready to use in class. Preliminary testing was carried out concerning its technical and educational issues as well as user satisfaction.

In addition, it uses some important characteristics of AR to foster learning, namely, real time annotation, contextual visualization and vision-haptic visualization (SANTOS, CHEN, *et al.*, 2014). In the ARBlocks, visual and audio feedback can be inserted in the activities. The tactile sense is explored through the use of tangible interaction which is also an effective way to scaffold young children's learning.

Although still lacking an authoring tool, it is also important to highlight that the tool is designed to be flexible and allow users to create their own activities. This is an important aspect since one possible solution for integrating AR systems into regular school curricula according to (WU, LEE, *et al.*, 2013) is to support teachers to tailor AR technologies, to create personalized learning tasks and to monitor student's progress in AR.

5.1.3.1. Educational Analysis

In order to previously evaluate ARBlocks' educational potential, eight teachers from different private and public schools were invited to use the tool and answer a few questions about it. All teachers work with children and had no previous knowledge about AR (ROBERTO, 2012).

Questions investigated their opinion about the concept, if they liked the experience of using it, if they had difficulties using it, if they believed the tool has potential to improve their classroom, if they believed the children would like to use it, the positive and negative aspects about it and, finally, if they had any improvement suggestions.

The teachers highlighted that by using physical blocks, children would be able to manipulate them, thus, the subject may be more meaningful and attractive.

The authors also reported that teachers referred as a positive aspect that the ARBlocks is very interactive and enables feedback and content change during manipulation.

The multidisciplinary aspect and versatility was also highlighted since with the same set of blocks it is possible to create different activities. They also mentioned it will enable them to be creative and guide the activities according to their needs.

Teachers claimed that since the ARBlocks is a technological tool it can help to increase student's motivation and hold their attention. They highlighted that students are used to technology in their daily lives.

Authors emphasized that teachers did not need any training to use the tool. The ARBlocks, thus, stands out as being a simple and intuitive tool.

Some of the teachers suggested that the tool could be improved if both faces of the blocks could be used to project the content.

5.1.3.2. User Satisfaction Analysis

Preliminary tests concerning user satisfaction were carried out using the drawing intervention technique. Authors argue that this technique can extract more subjective information about children's perception of the system (ROBERTO, FREITAS, *et al.*, 2013). The technique consisted of asking students to make drawings related to the experience and, then, have a post-conversation with them about their work.

The ARBlocks was used for one month in a public school in Recife, Pernambuco, Brazil. After eight sessions using the application, thirteen children were randomly chosen and asked to draw anything they wanted related to the tool. Those children had no previous contact with AR.

Authors report that children interacted with all the setup of the tool, some calling it a robot. However, (ROBERTO, FREITAS, *et al.*, 2013) mention that they were able to have the post-conversation with only seven

children since the others did not draw content related to the tool.

Teacher's point of view was also investigated through observation and a semi-structured interview.

Results show that students preferred the ARBlocks rather than other school material available. (ROBERTO, FREITAS, *et al.*, 2013) report that all students wanted to be part of the first group to use the tool. They also were willing to take it home. Positive feedback was given regarding the audio feedback. Children used to 'sing along' with claps and whistles when they hear it.

Negative aspects were also observed. Children and teachers had to stand up around a table to use the tool. This was tiring especially to the teacher who needed to manage sessions with all groups. Audio feedback was sometimes difficult to hear due to external noise. Children seemed not to worry about small jitter, projector drift or partial occlusion. The teacher, on the other hand, seemed to be a bit worried that children would tumble or make reckless movements that would crash the system.

Finally, the teacher was interviewed about the experience of using the ARBlocks. She recognized to be worried about the tool's safety. She pointed out that the feeling decreased over time. Teacher explained that the process of taking children to the library, where the sessions happened, for every activity was both positive and negative, since this change stimulates them but also required time to get them focused to start the activity. She also evaluated the tool positively in the sense that it helped students to engage in the reading process. She pointed out that some inattentive students were focused on the activities. She also found the tool attractive to the children.

5.2. PILOT STUDY

Before the case study was applied, a pilot study that evaluates the use of the ARBlocks system using different metrics and methods combined to measure its impact in the learning process was carried out (SILVA, ROBERTO and TEICHRIEB, 2013). The usability aspects of the tool were also part of the evaluation as previously described in the section about the ARBlocks.

The goal of this pilot study was to evaluate the impact of the ARBlocks in the literacy progress of year one students.

5.2.1. THE SCHOOL

The school participating in the study is a small public school placed in Recife, Pernambuco, Brazil[7]. The school is placed in a poor region of the city and most of its students come from the community around it.

This school offers early childhood education and elementary school in the mornings and afternoons. At night, it also offers adult education.

The infrastructure of the school consists of: six classrooms, a library with approximately one thousand books, tv, dvd player and projector, a computer laboratory with approximately ten computers, a schoolyard, the principal's office, teachers room and secretary office.

[7] In respect to ethic issues, the name of the school, teachers and students will be omited. In addition, pictures and videos were carefully taken during the experiment in order to preserve subject's identity.

The school also participates in the *Mais educação* program [8], therefore, students have different activities such as dancing lessons, chess, judo, math and literacy.

5.2.2. PARTICIPANTS

The teacher involved in the study has a degree in pedagogy and specialized in children education. She has twenty-one years of experience teaching young children.

She teaches two year one classes, one in the morning and the other one in the afternoon shift.

5.2.3. THE YEAR ONE GROUPS

During elementary schools, students are expected to master a range of abilities that involve:

- Understand citizenship as social and politic participation;
- To take position in a critically, responsibly and constructively way in different social situations;
- Know essential characteristics of Brazil in the social, material and cultural dimensions;
- Know and value the different Brazilian social and cultural patrimony as well as social and cultural aspects from other people;
- See yourself as part and transforming agent of the environment;
- Develop self-consciousness;
- Know and take care of one's own body and health;
- Use different languages – verbal, mathematical, graphic, plastic and body language as a way to express and communicate ideas as well as understand and enjoy cultural productions;
- Know different technologic and information resources;
- Inquire reality, formulate and solve problems.

During the year one, students work in part with all the aforementioned concepts as steps to master them at the end of the fourth year.

As regards to language which was worked in our experiment, they are expected to:

- Understand oral and written messages and begin to identify relevant information according to author's intentions;
- Read texts from genres established for the year (cycle), combining some basic reading strategies;
- Use oral language efficiently;
- Participate in different oral communicative situations;
- Produce cohesive and coherent texts;
- Write texts from the genres established for the year (cycle);
- Consider the different versions that a written text needs.

The book used in the school is (BRAGANÇA and CARPANEDA, 2008). The teacher highlights that she uses different sources to complement the material. The case group, which used the ARBlocks during their lessons, has approximately twenty students who have lessons in the morning shift. The control group also has approximately twenty

[8] The *Mais educação* program is a strategy of Brazilian education ministry to induce the extension of school day and curricular organization in the full-time education perspective. Schools enrolled in the program opt to develop activities for teaching support, such as environmental education, sports and leisure, human rights in education, culture and arts, digital literacy, health promotion, communication, media usage, field investigation of science and economics [110].

students, however, these students have lessons in the afternoon shift.

5.2.4. Study Design

In order to evaluate the impact of the ARBlocks in the literacy progress of year one students (age range 6-7 years old), a semi-experiment design was employed. Two quantitative metrics were used.

Two year one groups of a single teacher were selected for the research. One of them has classes in the morning shift and the other one in the afternoon shift.

Each class has approximately 20 students. In the beginning of the year, she applied a test with all her students in order to find out their literacy stage.

Based on the results of this test, we decided, along with the teacher, to apply the tool in the morning shift group, which presented the lowest scores. The afternoon shift was our control group. The tool was applied in the morning shift group twice a week for four weeks.

The first one was a written test at the end of the use of the ARBlocks. This test was designed to measure student's ability to recall the content studied with the tool. Although most of the evaluation studies use this type of test, it has some limitations. It is widely known the difficulty of measuring learning with a single test.

Studies have shown the importance of using a formative assessment in which students' skills are continuously evaluated throughout the learning process (ZABALA and DA ROSA, 2007).

In order to complement our metric, the teacher's own formative evaluation was used as a second quantitative metric. It provided a better overview of the student's progress through the year.

Her evaluation assesses student's literacy skills. It is based on the psychogenesis theory of written language learning developed by (FERREIRO, 1985).

The teacher at the beginning of each year applies an activity with her students to identify the literacy stage they are in according to this theory. She repeats the same test every two months in order to assess student's progress and current needs regarding literacy development.

The stages of written development proposed by (FERREIRO, 1985) are the following:

- Pre-syllabic 1: learners do not understand the relationship between oral and written language. They are able to write using drawings, scribbles or wavy lines;
- Pre-syllabic 2: learners can trace letters although they do not understand the correspondence between written and oral language;
- Syllabic (quantitative): learners write one letter per syllable;
- Syllabic (qualitative): learners use one letter per syllable, however, they try to use letters that are related to what they hear;
- Syllabic - alphabetic: learners are able to establish a relationship between graphemes and phonemes in most words although they still write units smaller than a syllable;
- Alphabetic: in this stage, learners can establish the relationship between letters, words and syllables.

As part of the qualitative metric, it was conducted a semi-structured interview with

the teacher in order to find out her perception of the children development and the use of the ARBlocks.

(FITZPATRICK, 2004) stress the need to involve teachers in the process of adopting new technology so the activities are integrated to their lesson plan and meaningful to the students.

In order to involve teachers in the adoption of technology, we used the flexibility of the ARBlocks to support different applications and encouraged the teacher to create the applications she needed.

Since the system does not have an authorship tool, the teacher needed to describe the activities to a programmer who developed them.

5.2.5. Environment System and Setup

The experiments were conducted in the school's library since it provided a good space for the system setup and the students. The researchers provided the equipment in order to facilitate assembly and to provide some equipment that were not available in the school. The ARBlocks run in an ordinary laptop having an Intel Core 2 Duo with 2Ghz, 4 GB of RAM, an integrated graphics card, a built-in speaker for the sonorous feedback and Windows 7. The computer was connected to an Epson projector EB-X10, similar to those found in several schools nowadays. It was used a Microsoft webcam LifeCam Cinema, that is also a standard model.

The projector was attached to the Artograph Digital Art Projector Tripod and pointed down to one of the library's table. The webcam was taped on the top of the projector in order to see the entire projection area. Figure 28 shows the environment and the system setup used in the experiments.

Figure 28: ARBlocks' setup

5.2.6. Activities

The activities proposed by the teacher involved mostly reading skills and phonemic awareness since this is one of the first steps to reach reading competence (ADAMS and FOORMAN, 2005). She requested four activities in total. The sections were organized so that each child could interact with the blocks every visit. The class was divided in three groups of six to seven participants each, depending on the number of students attending the class. While one group was using the system, the others were engaged in other educational activities related to the same topic. Each session lasted about 30 to 45 minutes depending on the activity.

One of the activities requested students to match pictures (e.g.: rat with cat). After doing the correct matching the students listened to the rhyme.

The next activity was based on matching but instead of regular pictures, students had to match their own drawings with situations that rhymed with their names (e.g.: Peter is

using his computer) as can be seen in Figure 29.

Figure 29: Activity in which students had to match their own drawings with situations that rhymed with their names

The last two activities requested involved nursery rhymes. In these activities, students were required to complete a nursery rhyme they had been studying in class. The nursery rhyme appeared on the table with some words missing that were displayed in the blocks. Students should place the words in order to listen to their own voices singing the nursery rhyme as shown in Figure 24.

5.2.7. EVALUATION

After the period using the ARBlocks, a test was applied with the morning shift group that used the tool and the afternoon one that did not use it, which was the control group. The test aimed to verify the content retention. It was elaborated by the teacher so students were familiar with it. The test consisted of four questions[9], three involving the content studied with the ARBlocks and one involving contents that were not practiced using the tool.

As previously mentioned, for the purpose of our evaluation, we considered the assessment the teacher does periodically with the students. Her evaluation consists of presenting some pictures to the students and asking them to write their names the way they believe they are written. After that, they read what they wrote to the teacher so she can classify them according to Ferreiro's stages of written development. It is the same test she applied in the beginning of the year.

As a qualitative evaluation, we conducted a semi-structured interview to understand the teacher's perception of the process of using the ARBlocks in her lessons. The interview was conducted in Portuguese and the questions asked were the following:

1. Quais os pontos positivos de utilizar a ferramenta na sua classe? (What were the positive and negative aspects of using the tool in your classroom?)
2. Quais os pontos negativos relativos ao uso da ferramenta? (What were the negative aspects of using the tool?)
3. Numa escala de 0 a 10, que nota você daria a ferramenta? Justifique. (In a scale ranging from 0 to 10, what score would you give to the tool? Explain.)
4. Que sugestões você daria para a melhoria da ferramenta? (What suggestions would you give to improve the tool?)
5. Você mencionaria alguma contribuição da ferramenta para o aprendizado de seu grupo-classe? (Would you mention any contribution of the tool to your learning group?)
6. Qual a probabilidade de você planejar suas aulas utilizando essa ferramenta? (What are the chances

[9] In apppendix D, the final test applied in the pilot study is presented.

for you to plan your lessons using this tool?)

7. Você tem alguma preocupação em relação ao uso dessa ferramenta? (Do you have any worries concerning the use of the tool?)

The conversation was audio recorded and lasted approximately 20 minutes.

5.2.8. Results

Following the evaluation methodology adopted in this work, the presentation of the results obtained will be divided in three parts: results of the final test applied after students used the tool, results of the teacher assessment and, finally, results of the interview conducted with the teacher.

5.2.8.1. Final Test Results

The final test applied after the period using the tool to both groups showed that the morning group had an equivalent score compared to the afternoon one regarding the questions about the topics worked with the ARBlocks.

The chart seen in Figure 30 presents the average score for both groups divided by questions using grades from 0 to 10. The morning group, which was in a lower stage of writing development in comparison to the afternoon group, achieved the same grade for the first question and a slight beneath score for the second one (7.88% lower). Both questions were addressed to rhyming.

In the third question, about the nursery rhymes, the morning group achieved a slightly higher grade than the afternoon shift group (5.67% higher).

In the last question, related to filling in the missing letter in the words, a subject that was not worked using the ARBlocks in both groups, the difference on the score was more evident (23.69% lower).

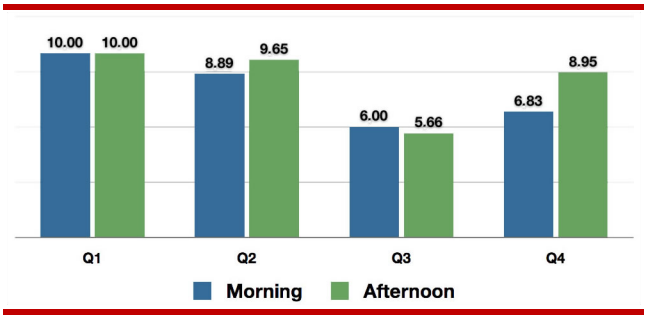


Figure 30: Average score of the morning and the afternoon shift groups separated by questions. Questions 1 to 3 were about topics worked with the ARBlocks and question 4 was taught without the help of the tool

5.2.8.2. Academic Performance Results

The second quantitative metric was the formative assessment the teacher regularly applies with the students throughout the year. The chart in Figure 31 reveals that in February, 2012 most of the morning students were in the pre-syllabic 2 stage. In April, 2012 the students were in both syllabic stages. Between April and June, 2012, none of the students could be found in a more advanced writing stage, but they were all established in three stages (pre-syllabic 2 and syllabic quantitative and qualitative).

Between June and September, 2012 students were migrating through these three stages. The ARBlocks was used with the morning group during four weeks between August, 2012 and September, 2012. The chart shows that after this period most of the students were in the syllabic stages and only four progressed to more advanced levels. Nevertheless, the teacher made this evaluation two weeks earlier in comparison to the two previous assessments.

In December, the majority of the class was in the syllabic alphabetic stage. The

rest of the class was spread between the syllabic stages.

In Figure 31 we also present the results of the teacher's assessment with the afternoon group. In the first assessment (February, 2012) we can see that most of the students were in the pre-syllabic 2 stage. However, there were students in all the stages of the development. In April, students started migrating from the pre-syllabic stages to the qualitative stages, although the majority of the class was in the pre-syllabic 2 stage. The syllabic-alphabetic and alphabetic stages did not change.

In June, more students reached the syllabic-alphabetic and alphabetic stages. In September, the majority of the class was in the syllabic qualitative stage. By the end of the year, most of the students were spread over the three final stages, syllabic qualitative, syllabic-alphabetic and alphabetic. Four students remained in the pre-syllabic 2 stage.

5.2.8.3. Interview Results

Finally, in the semi-structured interview, the teacher was questioned about the use of the ARBlocks. According to her, the main advantage of the ARBlocks is that it provides different and enjoyable playful activities that can be related with the content worked in the classroom. She mentioned that the school has several educational software, but when the students go to the informatics laboratory they often use applications that are not linked with the classroom content.

The teacher also mentioned that students were more willing to read during the sessions and exemplified that one of the children started reading during the one activity with the tool.

She also noted that their advancement was faster than expected and that parents were noticing student's progress and commented with her.

5.2.9. Discussion and Further Improvements

The results of the final test applied after the use of the tool showed that although the morning group started the year in a lower level of development, they reached similar levels to the afternoon group. In the last question of the test that was worked in the traditional way in both groups, the difference of their scores was higher. This suggests that AR seems to have helped them to better grasp the content.

The students seemed to be very motivated by the tool and its use. They appeared to get attached to the tool and named it "the robot". The sonorous feedback was perceived as the ability of the robot to speak. They seemed to be excited in engaging in an activity different from their routine. These observations led the authors to discuss about how to design educational AR applications (ROBERTO, FREITAS, *et al.*, 2013).

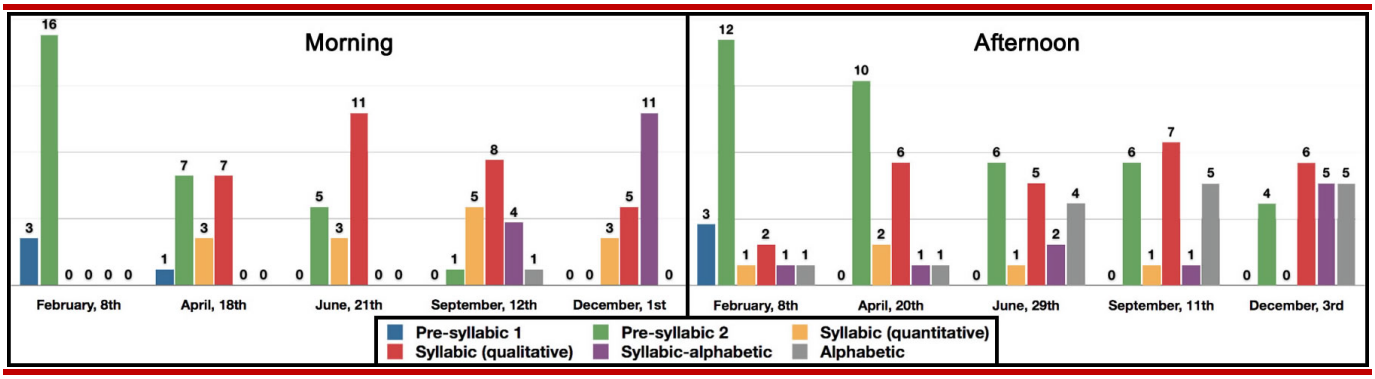


Figure 31: Number of students from the morning and afternoon shifts according to their writing development in the teacher's evaluations through the year

During the interview, the teacher mentioned that children enjoyed using the ARBlocks and this extra motivation made them more focused to read. She said that the morning shift group did not seem to be willing to read and a few days after starting using the tool they became more interested in doing the activity.

The teacher claimed that the ARBlocks played an important role in children's educational progress. According to her, it was not expected that any of the morning students would be in the alphabetic stage by September and still there was one pupil in this stage and two more almost there.

In her opinion, the progress was faster than expected. She also mentioned one student that started to read during one of the ARBlocks' session. The teacher believed that this happened because this child became very motivated after using the tool.

It is important to mention that the student who reached the alphabetic stage during the use of the tool finished the year in the syllabic-alphabetic stage. This is normal since the written development is not a linear process. Learners are always testing and retesting their hypothesis about the written language and, thus, consolidating their knowledge.

Although preliminary, the results supported the hypothesis that the ARBlocks can help to motivate students and foster the development of their literacy skills. The teacher provided a positive feedback regarding the tool and its use, highlighting its flexibility.

Regarding the teacher's regular evaluation, we believe that it is a good way to have an overview of student's development during the year. This is important since it's a formative evaluation, which may represent better student's development rather than just a punctual test. Through the study, we could notice that in a smaller period of time, the children reached a good progress concerning their writing development.

In the test applied after the ARBlocks use, we could see that in the questions that had been worked with the tool, students reached a satisfactory score.

The interview of the teacher along with the researcher's observations were also valuable resources for better understanding the process of introduction and impact of the AR tool in this learning context. Qualitative methods are usually associated to social sciences and qualitative research usually focuses on detecting and processing intentions (REDONDO, VALLS, *et al.*, 2014). Qualitative methods are emphasized

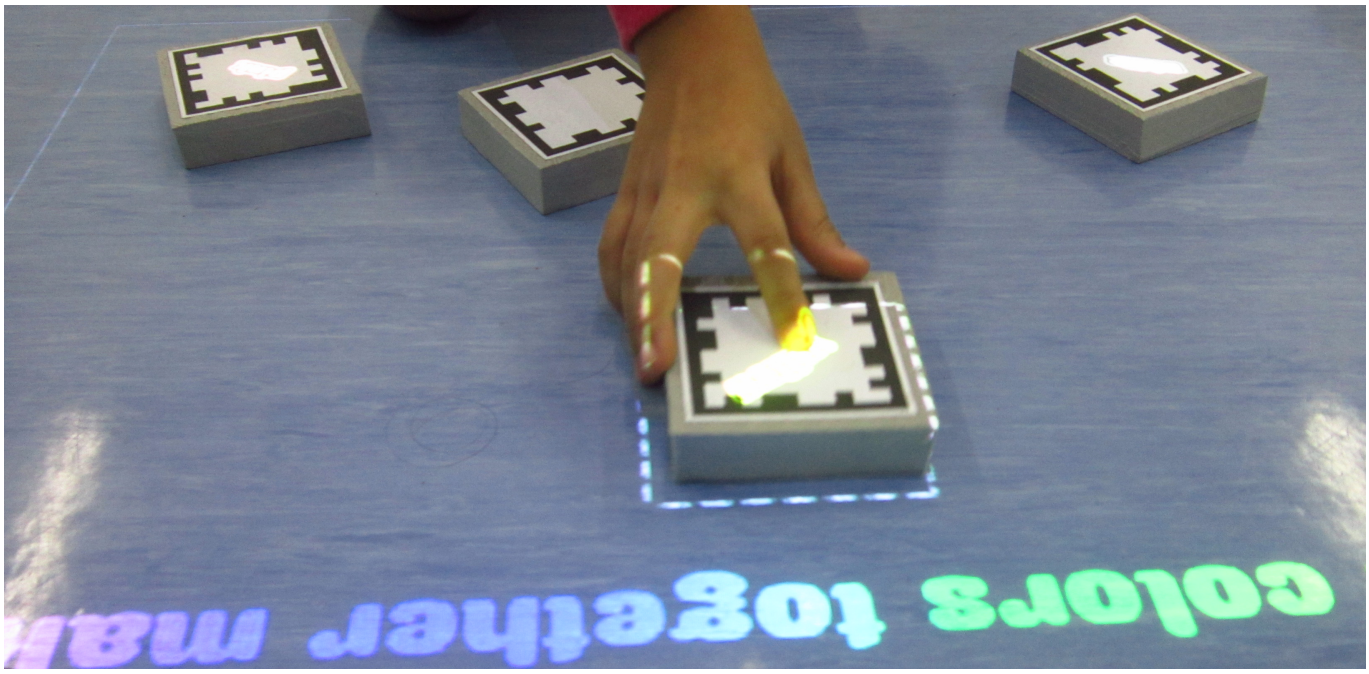
especially with young children (MOONEY, 2013). In this sense, teacher's observations and comments are very valuable since they play an important role in the children's development, especially, concerning the acquisition of formal knowledge in schools. Through the interview, the teacher seemed surprised to see one of her students in the alphabetic stage in September. We noticed that she was very enthusiastic with the use of the tool and she believed that the students were extremely engaged in the activities.

For further works, we believe that the system must be used for a longer period of time and with more groups of students. It is also important to test the tool with different teachers in order to discern and evaluate the impact of teacher's methodology in the use of the device. These approaches might help to hinder the confounding aspects that may affect the evaluation. Additionally, statistic measures can be applied to the data in order to reinforce the conclusions obtained in the qualitative tests.

The evaluation using different metrics should be encouraged in further studies since it facilitates to have a better overview of the impact of the tools along with teacher's involvement in the process of introducing the tool and evaluating its impact and potential.

6. EXPERIMENTAL STUDY

Figure 32: Pre-Kinder 2 students using the ARBlocks to practice color mixing vocabulary



After the pilot study and the lessons learned through the experience, a quasi-experimental study was carried out in order to further investigate the impact of AR technology in the educational field. We designed a quasi- experiment (EASTERBROOK, SINGER, *et al.*, 2008) in order to examine the impact of the ARBlocks, a projective AR tool especially designed for children education in the English language teaching (SILVA, ROBERTO and TEICHRIEB, 2015).

As previously mentioned, the system chosen to be evaluated was the ARBlocks. This tool enables teachers to create the activities they will use. Hence, the teachers previously met the researchers in order to discuss the activities they needed. The activities requested were implemented by the researchers following teacher's specifications.

The specific objectives of our study are:

1. to evaluate the use of the ARBlocks in the English teaching field concerning the following aspects: linguistic concepts and competencies;
2. discuss possibilities of use for AR tools in the classroom environment as well as reflect about the potential and difficulties involved in the introduction of this new piece of technology in the school environment;
3. provide some guidelines in order to assist researchers when conducting similar evaluations.

In order to achieve the aforementioned objectives, it was designed the quasi-experiment study. Researchers will make sure that each teacher involved in the study will have access to two different contexts: use and not use of the AR system.

It were used different metrics in order to compensate the weakness of each approach. Two quantitative metrics were used: the academic performance of the students and a final test after the use of the system in order to check student's understanding of the content. Two qualitative metrics were also used: a semi-structured interview with the teacher in the end of the entire process and a research diary which teachers filled in after every session. Tape recordings and observations were also used in order to have a better overview of the process.

In addition to that, a questionnaire was applied with the students to evaluate their satisfaction of using the ARBlocks as well as understand their perception of the tool and its use. A diagram with the entire process can be seen in Figure 33.

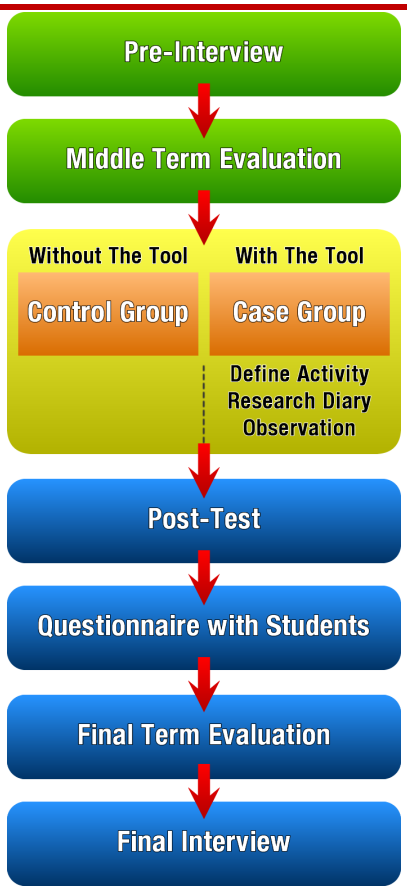


Figure 33: Diagram with the methodology design of the study. The middle term evaluation and the final term evaluation correspond to the academic achievement of the students

The tool was used in a language school[10] in Recife-PE, Brazil, for approximately three months for two different groups of young students learning English as a foreign language. Age of students ranged from 6-9 years old. The groups were in two different levels, namely: Pre-kinder 2 (6-7 years old) and Kids 1 (8-9 years old). Each group was taught by a different teacher (described in this study as teachers A and B, respectively). Two groups of each grade level participated in the study. One group of each level used the tool (case groups) and the others did not use the ARBlocks in their lessons (control groups).

It is also important to highlight that since it is a language course the curricula is designed for a semester which provides a better overview of the process since it was not possible to spend the whole year in the experiment. In addition, as will be latter discussed the school chosen is known for its interest in applying technology tools in its lessons. Hence, it is important to compare the use of the ARBlocks with different technology tools since traditional paper and pencils are usually not very motivating for students.

6.1. The School[11]

The language school participating in the study is one of the biggest English Language schools in Recife, Pernambuco, Brazil. It is

[10] In respect to ethic issues, the name of the school, teachers and students will be omited. In addition, pictures and videos were carefully taken during the experiment in order to preserve subject's identity.

[11] Information in this chapter were provided by the teacher, coordinators and researcher's observation. School coordinators were allowed to check this final report before it has been published.

placed in a wealthy and urban area of the city.

The school was chosen for its known interest in applying new technologies in the English teaching process. It has a smart board in each class and students also have at their disposal the following infrastructure:

- Computer laboratory (Cybers);
- iPad's mobile laboratory;
- Library with more than ten thousand items divided among books, DVDs and pedagogic games;
- Kids space with computers and video-games, such as the XBOX;
- 3D rooms;
- Auditorium;
- Recording studio;
- Media Lab equipped with Mac OS computers for video and audio projects, mostly used for older students.

Teachers are also constantly trained in order to foster their teaching skills and language abilities as well as discuss ways of using new technology adequately in class, new trends in language teaching and organizational and operational issues (e.g. teacher-parents relationships, grading system).

Once a month, teachers get together in order to discuss teaching trends. Teachers explain that the training sessions usually consist of colleagues sharing experiences related to online courses taken from Harvard University. They have the opportunity to reflect about those trends and their relationships with their own experience. They are encouraged to adapt their lessons when necessary.

The school also promotes some projects every semester for students, such as the poetry month[12] in the first semester and the banned books[13] in the second one. Each level has also its own projects, such as movie award[14] in which they record a movie and present it to the audience and stop motion projects in which students use stop motion techniques to craft animations.

The school curriculum and coursebooks of the school have been recently adapted to encompass the twenty-first century abilities which are worked throughout the entire course. Those abilities are divided in four groups as described above:

1. thinking skills - creativity, critical thinking, problem-solving and decision-making skills;
2. working skills – communication and collaboration;
3. skills to live in the world – citizenship, carrier preparation, social and personal responsibility;
4. working tools – information technology use for communication, comprehension and information.

6.2. THE TEACHERS

In this section, it will be provided some information about the teachers involved in the study.

6.2.1. TEACHER A

The teacher of the Pre-Kinder 2 groups is undergraduated in language teaching (i.e.,

[12] The poetry month is a project dedicated to explore literature in English language teaching. The library is the place for most of the activities related to the project. Poems are exposed to students from all levels in the school. Teachers and students read and do some activities related to the poems.

[13] The banned books project explores banned books and the power of writing. The library is the place for most of the activities related to the project. Students can discuss why those books were banned. Young children can participate in storytelling, reflection and craft making activities.

[14] This project usually occurs with older students.

Portuguese and English). During her undergraduation, she had experience with scientific research. She has the ECPE certificate (Examination for the Certificate of Proficiency in English) from Michigan University and has been teaching English to different levels for seven years. She teaches two Pre-kinder 2 groups in the current term.

6.2.2. TEACHER B

Teacher B is undergraduated in Computer Science. She has different English language certificates, i.e., SAT (Scholastic Aptitude Test), ECPE from Michigan University and TOEFL (Test of English as a Foreign Language). She has twenty-three years of experience teaching English. She teaches two Kids 1 groups in the current term.

6.3. THE GROUPS

In this section, it will be provided some information about the groups involved in the study.

6.3.1. THE PRE-KINDER 2 GROUPS

The language expectations for this group are all related to children's reality. At this level, children are already exposed to grammar structure although it happens naturally and gradually. Children need to be able to read, write, speak and listen to vocabulary related to the following topics:

- Greetings;
- Toys;
- Farm animals;
- Zoo animals;
- Colors;
- Mix of colors;
- Numbers from 1-20;
- Math operations: + and -;
- Parts and objects of the house;

-Prepositions of places using parts and objects of the house;

- Shapes;
- Nature.

The twenty-first century competencies are also emphasized during the program. The competencies explored at this level are: communication, critical thinking, creativity, collaboration and caring.

The book used is the Bounce Now (KNIVETON and LLANAS, 2011). The book is divided in nine units related to the topics mentioned above. It is important to highlight that throughout the units, the topics are constantly reviewed. All units present activities related to the four language skills. At the end of each unit, there are final activities. All of these activities explore writing skills.

Teacher explains that she explores the contents using not only the book but different resources that students enjoy, such as the smart board, websites, videos and games. She points out that she always shifts from one activity with movement to another one that they need to concentrate on. They start every class with homework correction. In the end of the semester, she reinforced writing and reading skills since most of the students are still learning them at school and, therefore, present some difficulties.

Students are evaluated holistically. They are evaluated every class and they also have a reflection time. Students at this level do not work with lots of worksheets. In two different moments of the semester, they get together with the teacher in order to reflect about what they have learned so far and what they still have to concentrate on to improve.

6.3.1.1. Case Group

The Pre-Kinder 2 case group has lessons in the morning shift twice a week. Each class lasts one hour and fifteen minutes. It is composed by twelve students ranging from six to seven years old. There are seven boys and five girls.

The group is dynamic. They enjoy musical videos and funny sounds. They love group activities especially competitions. The teacher explained that movement is essential to motivate them to do the activities, however, they need to do many activities involving writing or drawing that require them to be seated as well.

6.3.1.2. Control Group

The Pre-Kinder 2 control group has lessons in the morning shift twice a week. Each class lasts one hour and fifteen minutes. It is composed by ten students ranging from six to seven years old. There are five boys and five girls.

The group demonstrates behavior problems, especially, the boys. The teacher claims that she has to repeat instructions before applying the activities. They demonstrate a good development. Just one student presents difficulty in letter recognition.

6.3.2. The Kids 1 Groups

Kids 1 groups are expected to master vocabulary related to the following topics:

- Welcome to our class;
- My classroom;
- My world;
- Family;
- My house;
- Cool clothes;
- My toys;
- My body;
- Good food;
- Animal friends.

The twenty-first century competencies are also emphasized during the program. The competencies explored at this level are: communication, critical thinking, creativity, collaboration and caring.

The book used is Our World (PINKLEY, 2014). The book is organized in nine units related to the aforementioned topics. The units work with the four language skills and always relate the topic at hand to different cultural habits around the world.

Each unit ends with a small project, such as making a poster or a robot. At the end of tree units, the books always propose a game to review the content seen so far.

The teacher explains that Kids 1 classes are very dynamic. The teacher also points out that she always shifts from one activity with movement to another one that they need to concentrate on. She also adds that they explore all the abilities as much as possible. They work a lot with speaking and listening skills. At the end of each unit, students write simple sentences about the content they have just learned. They also get involved in small projects, such as writing a poster or a booklet and then present that to the group. Teacher emphasizes that they also use the iPad and Computer laboratories which are very appealing to students.

The teacher also explains that she uses lots of body language and that they have a routine every class involving the date, climate, day of the week, classroom language and commands. They also have homework correction every class.

Students are evaluated holistically. They are evaluated every class and they also build a portfolio. Students collect activities done

during the classes and in two different moments of the semester, they get together with the teacher in order to select important work and reflect about what they have learned so far and what they still need to concentrate on to improve.

6.3.2.1. CASE GROUP

The Kids 1 case group has lessons in the morning shift twice a week. Each class lasts one hour and fifteen minutes. This group has nine students, eight girls and one boy. These students present good behavior. However, most of the students are new in the school (i.e., it is their first year learning English), thus, they present a lower linguistic level. One of the students is repeating the level.

6.3.2.2. CONTROL GROUP

The Kids 1 control group has lessons in the afternoon shift twice a week. Each class lasts one hour and fifteen minutes. This group has eight students, seven boys and one girl. This group presents some behavior problems. Their linguistic level is good since most of the students were studying English in the school since the kinder stage (age range through 7 to 8 years old).

Teacher B (from Kids 1 group) decided to use the tool once a week.

6.4. EXPERIMENT DESIGN

In order to evaluate real class interaction it is important to establish a partnership with teachers and respect their particular routines. Before we started the experiment, researchers got together with the teachers and coordinators to explain the objectives of the research and present the tool to be evaluated, the ARBlocks. The teachers had the opportunity to manipulate the tool and see some applications. After this process, they participated in a semi-structured interview about the tool and its potential. The interview was carried out in Portuguese. The questions asked were the following:

- Descreva suas impressões sobre a ferramenta (Describe your impressions about the tool)
- Você consegue enxergar potencial desta aplicação em sua sala de aula? Explique. (Can you see potential of this application in your classroom? Explain.)
- Que tipos de conteúdo poderiam ser ensinados com a ferramenta? Quais habilidades linguísticas poderiam ser ensinadas com a ferramenta? (What kind of content can be thought with the tool? What language abilities can be thought with the tool?)

Researcher explained to the teachers that they were able to generate ideas for their own activities. They decided the activities and explained that to researchers. Researchers were responsible to make the activities the way teachers intended.

The teachers, coordinators and researchers decided to choose which groups would be the case and control groups.

Teachers decided how to use the tool as long as it would fit their classroom needs. The tool was used in the Pre-Kinder 2 group twice a week (i.e., every class).

Teachers received a research diary[15] with some questions to fill in about every section of ARBlocks' use. This resource was used since it enables researchers to get information about day to day activities and explore those information in a subsequent interview (JACELON and IMPERIO, 2005) (ORTLIPP, 2008).

[15] Check Appendix E.

Different metrics were used in the evaluation. First of all, it was considered teacher's own evaluation. All groups at this school have two evaluation moments at the semester, known as middle-term evaluation and final term evaluation. This was also a good indicator of students' previous knowledge levels. Students are evaluated holistically, concerning some aspects: (a) continuous learning (reflection and requirements), (b) behavior, (c) homework – frequency (i.e., if students do their homework frequently), (d) homework – performance (i.e., if students do their homework accurately), (e) participation and (f) speaking.

In the continuous learning aspect, teachers observe how is the student overall development concerning language aspects. They base their evaluation on every day observation and student's portfolio.

Children's behavior is observed as part of their evaluation. Aspects such as responsibility and self-reflection are also taken into account.

In the homework frequency aspect, it is observed if students bring their homework frequently. In the performance aspect, it is observed the accuracy of students in the homework activity.

As participation, teachers observe students' availability in taking part of the activities, their ability to cooperate in group work among other aspects.

For speaking, teachers observe student's everyday use of the language, their pronunciation and how much they learn and use new vocabulary studied.

The grades are given as concepts ranging from ED to DM as described in the table below.

Evaluation Concepts[16]	
ED	Excelente desempenho (excellent performance)
BD	Bom desempenho (good performance)
DS	Desempenho suficiente (borderline performance)
DM	Desempenho não – satisfatório (unsatisfactory performance)

Table 12: Evaluation concepts

The statistical analysis was calculated for both case and control groups concerning the middle term and final term evaluations applied by the school and the final test applied after the period using the ARBlocks. The analysis was perfomed on the spss program version 20.0. The significance level was 95%. The first statistical test used was the Kolmorov-Smirnov in order to check data distribution as well as categorize the variables as parametric or non-parametric. For the comparisons of the same group, the mean comparison of the parametric data should be done through the Paired T- test and for the non-parametric data the Wilcoxon test should be used.

For comparisons among the groups, the mean comparison of parametric data should be done with the T-test independent sample. For the non-parametric data, the Mann-Whitney test was used.

For the statistical comparison, the scores of the tests were converted into grades using the following scale provided by the teachers: ED = 9 - 10 BD= 8 - 8,9 DS= 7 - 7,9 DM = less than 7

[16] The translation of the concepts was made by the researcher.

Students started to use the ARBlocks after the middle-term evaluation and worked with the tool until their final term evaluation. Pictures and videos were taken during the sessions with extra care in order to preserve subject's identities.

In addition to that, at the end of all the sessions with the ARBlocks, students from all groups answered an activity in order to evaluate students learning of the content seen throughout the semester [17] . Their performance in the content seen with the tool was compared to their performance in contents not worked with the ARBlocks. The performance variation among all the groups participating in the experiment was also compared.

Students were also required to reflect critically about their experience using the ARBlocks. A questionnaire, similar to the ones they use for self-evaluation during the semester, was provided [18] . They were expected to color the answer that best fits their opinion according to a smiley code.

Teachers also participated in a semi-structured interview at the end of the semester in order to evaluate the introduction of the AR tool in their classrooms. This interview was conducted in Portuguese. The questions in the interview were the following:

1. Como você analisa a experiência de utilizar o ARBlocks em sua sala de aula? (How do you analyze the experience of using the ARBlocks in your classroom?)
2. Como se deu o seu planejamento para o uso? (How did you plan for the use of ARBlocks?)
3. Houve dificuldades em relação ao planejamento? Quais? (Did you have any difficulties for planning? What were they?)
4. Como foi a interação com a ferramenta? (How was the interaction with the tool?)
5. Quais critérios você utilizou para escolher a dinâmica das atividades com os alunos? (What criteria did you use to choose the activities' dynamics used with your students?)
6. Como você avalia a interação dos alunos com a ferramenta? (How do you evaluate the student's interactions with the tool?)
7. Você considera que o uso do ARBlocks teve impacto na aprendizagem de seus alunos? De que forma? Se possível, mencione exemplos. (Do you think the use of the ARBlocks had any impact on student's learning? How? If possible, mention some examples.)
8. Algum comentário adicional sobre o uso da ferramenta que considere importante e não foi mencionado durante a entrevista? (Do you have any additional comments about the use of the tool that you consider important and was not mentioned during the interview?)

Teachers were also required to fill in a research diary about their everyday use of the blocks. This research diary intends to capture teacher's impressions of the use right after they have worked with the ARBlocks and, therefore, avoid forgetfulness. Questions cover many aspects such as the kind of activity used, skills practiced during the work, type of interaction with the tool and

[17] In apppendixes F and G, the final tests applied in the Pre-Kinder 2 and Kids 1 groups, respectively, are presented.
[18] In apppendix H, the questionnaire designed to acess students' is presented.

among teachers and students, students reactions and teachers impressions.

6.5. System Setup

The experiments were conducted in a special room provided by the school since it has appropriate illumination for the projection and enough space for the system setup and the students. The researchers provided the equipment in order to facilitate assembly.

In this experiment, the ARBlocks run in an ordinary laptop having an Intel Core 2 Duo with 2Ghz, 4 GB of RAM, an integrated graphics card, a built-in speaker for the sonorous feedback and Windows 7. The computer was connected to an Epson projector EB-X10, similar to those found in several schools nowadays. It was used a Microsoft webcam LifeCam Cinema, that is also a standard model.

The projector was attached to the Artograph Digital Art Projector Tripod and pointed down to the floor. The webcam was taped on the top of the projector in order to see the entire projection area.

6.6. Activities

The activities for the ARBlocks were all proposed by the teachers as they are the experts at their group needs as exemplified in Figure 32. The teachers acted as partners in the process of crafting activities specifically designed for their classrooms needs and school curriculum.

As previously mentioned, researchers got together with the teachers and coordinators in order to present the tool and show some previous activities as examples for them. The teachers and coordinators were allowed to explore the ARBlocks and understand its possibilities.

Teachers decided the activities beforehand and explained them to researchers in order to have them implemented for classroom use.

7. RESULTS AND DISCUSSION

Figure 34: Pre-Kinder 2 students using the ARBlocks to practice unscrambling words

The aim of our quasi-experimental study was to investigate the impact of AR technology in the educational field. Therefore, the quasi-experiment was designed in order to examine the impact of the ARBlocks, a projective AR tool especially designed for children education, in the field of English Language Teaching as a Foreign Language.

The study consisted of a quasi-experiment with two different groups of young students learning English as a foreign language. The groups were in two different levels, as previously exposed, the Pre-Kinder 2 (6-7 years old) and Kids 1 level (8-9 years old).

In this chapter, the results of the experiment will be properly described and discussed.

7.1. PRE-KINDER 2 RESULTS

In this section, it will be presented the results from the Pre-Kinder 2 groups.

7.1.1. PRELIMINARY INTERVIEW

In the first interview, the Pre-Kinder 2 teacher was asked about her impressions of the tool. The interview was audio-recorded and lasted approximately five minutes. The teacher thought the ARBlocks was excellent and very coherent with student's reality since it enables them to touch and listen to sounds. She explained that pupils liked to touch things and that their toys usually allow them to touch and listen to sounds. To sum up, she believed the tool was adequate for them.

When argued if she saw potential for the tool in the classroom, she was affirmative and mentioned that she already had six

ideas for different activities for her students. She believed that it is possible to work with the four language skills and mentioned that it would be similar to what she does in her classroom, that is, ask students to read, do things using their bodies, write and speak.

She also mentioned some possible activities for the tool, such as number sequence, number and vocabulary recognition, word order exercises and association of sounds to its correspondent pictures or names.

7.1.2. Activities

The activities proposed by the Pre-Kinder 2 teacher involved reviewing and reinforcing content seen in class through different sources. The teacher highlighted that the ARBlocks was a different way for students to review content and practice the reading skill which was important since the students were beginning to learn how to read.

She used the ARBlocks twice a week. Therefore, twelve activities were done with the students. The activities proposed involved reviewing and reinforcing content seen in class through different sources.

The first activity required students to answer to some questions by choosing the correct animal name (e.g.: tiger, lion, ostrich) as can be seen in Figure 35.

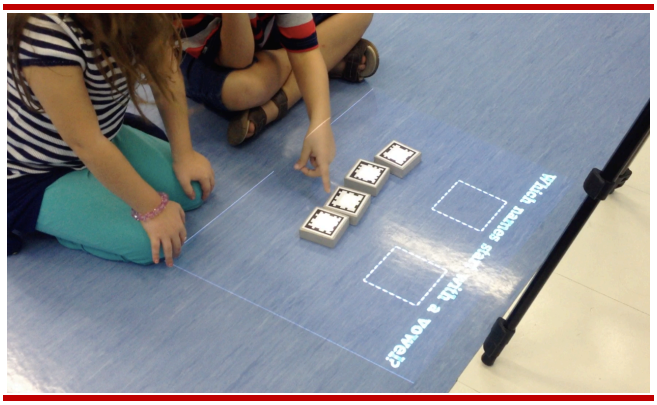


Figure 35: Students practicing vocabulary in the activity with the ARBlocks

Questions were related to different contents as can be seen bellow:

1. Which names start with a vowel?
2. Which names start with a consonant?
3. Which names have three letters?
4. Which names have four letters?
5. Which animals live in the forest?
6. Which animals are dangerous?
7. Which animals are pets?

In the second activity, students needed to match the numbers from one to ten with their respective names.

Another activity was based on gap filling. Students had to complete some sentences related to school material with the grammar structure “I have” and “I don’t have” according to the face that appeared next to the sentence (a happy or a sad face) as presented in Figure 36.

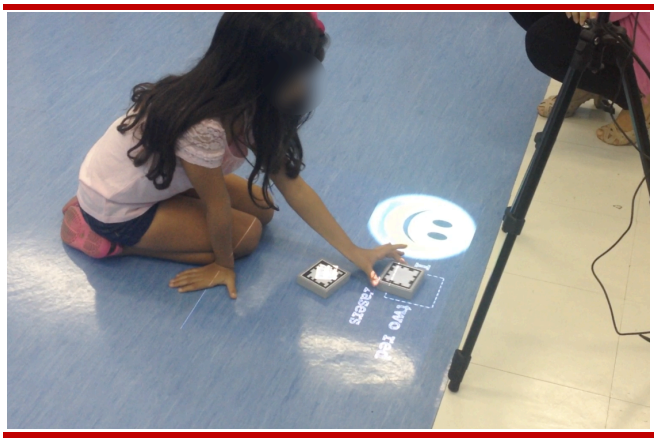


Figure 36: Student practicing the grammar structure I have and I don't have individually

The sentences were the following:

1. I ..... a blue pencil case.
2. I ..... a yellow pencil sharpener.
3. I .....a white school bag.
4. I .....two red erasers.
5. I .....ten crayons.
6. I .....one purple notebook.

In a different session, a quiz was proposed in which students answered questions about school material:

1. What school object do you use to cut?
2. What school object do you use to correct mistakes?
3. What school object do you use to stick pictures?
4. What school object do you use to draw?
5. What school object do you use to keep books, notebooks and pencil case?
6. What school object do you use to keep scissors, pencil, crayons and eraser?

The following activity was related to furniture. Some parts of the house were presented. Students had to associate the names of the furniture on the blocks to the correct room displayed on the floor as presented in Figure 37.

Figure 37: Students working with furniture vocabulary using the ARBlocks

Color mixing was the theme for another activity. Students were presented to a sentence (e.g.: green and yellow make…) and had to choose the correct color.

In another class, students had to make the names of the toys (e.g.: truck, ball, hoop) using the letters on the blocks as in Figure 34.

A different activity was proposed in which students needed to make the names of the colors (e.g.: blue, pink, red) using the letters on the blocks.

Another activity was related to sorting out words using “a” and “an”. Students were presented to different vocabulary seen in the classroom (e.g.: toys, transportation) and had to decide what article to use (a or an) before the names.

One of the activities was designed to practice geometric shapes. Students saw an image made of different shapes. They needed to count how many shapes were in the picture according to the question displayed (e.g.: how many hearts are in this picture?). Students chose the correct answer displayed in the blocks.

Another activity intended to review nature vocabulary. In this case, students saw an image of nature and had to answer some questions related to the picture (e.g.: what can you see?). Students needed to choose the correct answer.

Finally, the last activity aimed at revising number sequence. Students saw a sequence of numbers ranging from 1 to 20 and had to complete it with the missing numbers.

7.1.3. Research Diary

As previously mentioned, after every ARBlocks session the teachers were asked to fill in a research diary with some questions about the use of the tool. The teacher filled in the report for 11 sessions. She missed the report of one session.

To begin with, the teacher recorded some basic information such as the track, the time, the date and what was the activity done.

One of the questions was related to what were the skills practiced with the tool. Teacher mentioned reading, listening and speaking.

Regarding how the students interacted with the tool, students interacted individually (seven sessions), in pairs (four times)[19] and competition (one time).

The teacher interacted with the group during the activities in almost all sections, except in one of them. It is important to highlight that the teacher asked the researcher to conduct the activity two times even though she was interacting with the children.

She also reported that they interacted very well with the tool in all the sessions. For instance, she mentioned that “they loved the exercise. Everybody was anxious to take part of the activity”, “they were all enthusiastic to take part in the activity”, “they showed a very good comprehension about the rules and practiced very well” and “they showed confidence and pleasure in moving the blocks to make the name”.

When asked if she observed students learning something using the application, she mentioned some examples, such as: “they had to categorize the answers and associate the results with the listening”. “they used their previous knowledge of school material to understand the questions and face associated”, “review vocabulary, pronunciation and concepts”, “improve spelling of the names”, “they consolidated their knowledge about the colors mix”, “they learn especially to wait their turn”, “they understood they had to separate the blocks with names by observing if the names started with vowels and consonants”, “related shapes, numbers and previous contents”, “they focused on the listening skill”.

She did not mention any difficult of the students while interacting with the tool.

To conclude, she was asked to freely report any comment or situation. After the first section, she mentioned that they really loved it “once it is totally different from what they are used to”.

She also wrote some comments on the activity about color mix. She mentioned that “students were excited about the activity because they liked the theme and also because the activity had vibrant colors moving”.

She also commented in another session that it was nice that researcher divided them into two groups. The only problem was their behavior.

7.1.4. Academic Achievement

7.1.4.1. Case Group

Through the middle term evaluation, it is possible to observe that overall students presented a good development and command of the language.

In the first requirement, continuous learning, three students received ED concept and nine, BD.

The second requirement is the behavior, which is an important social ability that helps overall learning. Students were equally divided in three concepts, ED, BD and DS. Four students received each of these concepts.

Most of the students brought their homework frequently, therefore, ten children

[19] One of the sessions was not reported in the research diary by the teacher. In that session, she asked students to interact in pairs.

received the ED concept. Only two of them received BD for this concept.

Regarding homework performance, pupils were divided in two concepts, ED (six students) and BD (six students).

Regarding participation, four students received ED and eight received BD.

As regards to speaking, three students received ED, four of them BD and five, DS.

In the final term evaluation, students progressed in all aspects of learning.

Regarding continuous learning, five students were in the ED concept and seven in the BD.

In the behavior requirement, nine students received ED, one received BD and two DS.

In the homework frequency aspect, eleven students received ED and one BD.

Regarding homework performance, there was also improvement, eleven students received ED and one received BD.

Participation also increased, ten students received ED and two, BD.

In the speaking aspect, three students received ED and nine, BD. The results from both middle term and final term evaluations are properly displayed in Figure 38.

Concerning the statistic comparisons, firstly, the scores of the middle term and final term evaluation were compared with all groups together, namely Pre-Kinder 2 and Kids 1. In this case, there was no statistical significance. The final term evaluation revealed no statistical difference concerning these groups.

The statistical comparisons were also done by levels. In the middle-term evaluation, there was no difference in the scores of both case and control groups, therefore, the groups presented similar levels in all aspects evaluated.

The middle and final term results of the Pre-Kinder 2 case group were also compared statistically, as can be seen in Figure 39, most of the categories did not present significant difference. Only the homework-performance category presented statistical difference ($p=0.025$). This result may suggest that the use of the tool had a positive impact on student’s performance regarding the topics studied. As previously exposed, the other research instruments used reinforced this positive impact.

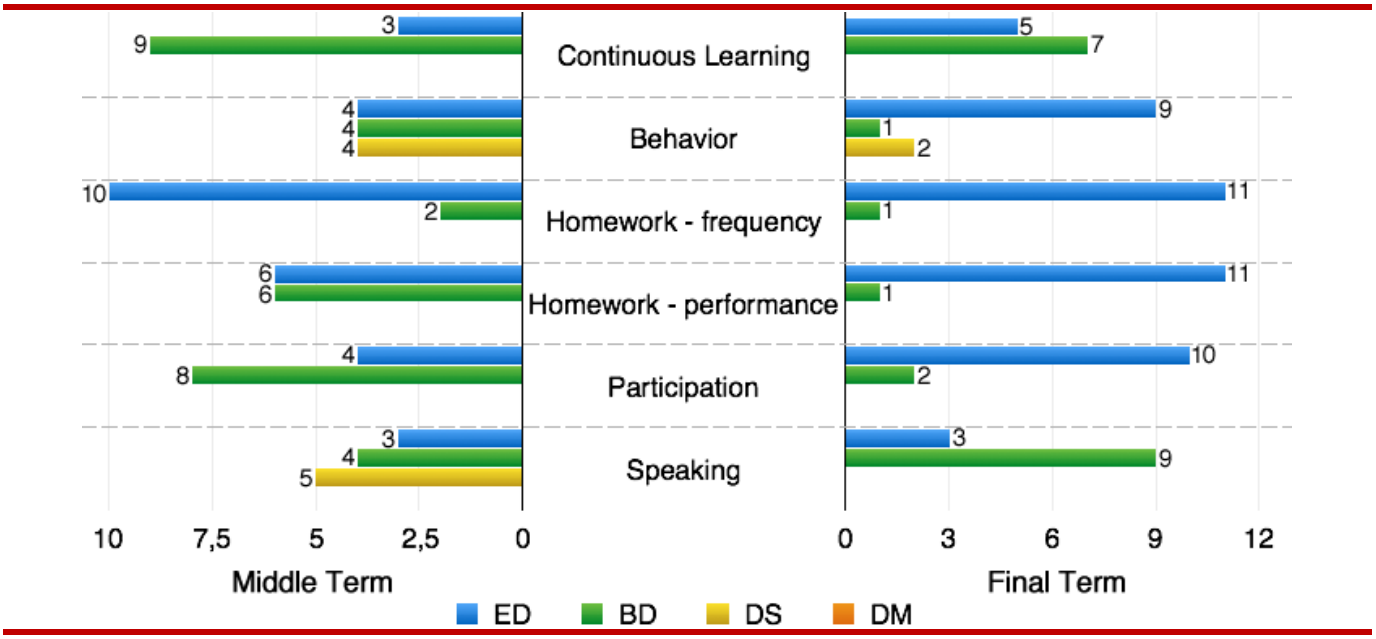


Figure 38: On the left side, Pre-Kinder 2 case student's performance in the middle term evaluation. On the right side, Pre-Kinder 2 case student's performance in the final term evaluation

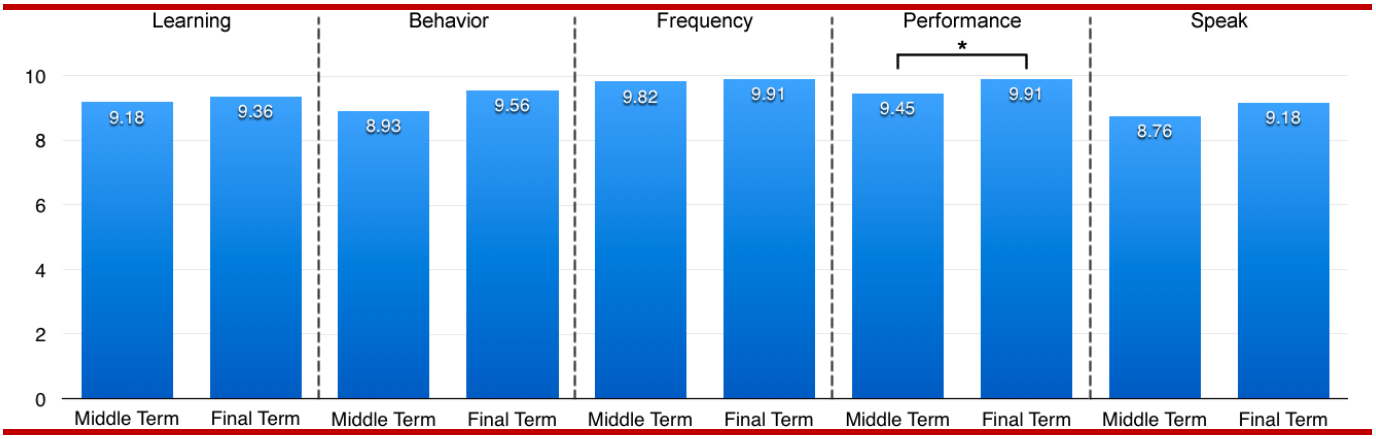


Figure 39: Pre-Kinder 2 case group results for comparison between middle term and final term evaluation. (*p<0.05 based on Wilcoxon test).

7.1.4.2. Control Group

The control group also presents a good development and command of the language.

In the middle term evaluation, six students received ED concerning continuous learning, and the other four students received BD.

As regards to behavior, four students received ED, four of them BD, one student, DS and one more DM.

Concerning homework frequency, seven students received ED and three BD.

Regarding homework performance, eight students received ED and two of them, BD.

In the participation aspect, three students received ED, six earned BD and one obtained DS.

With respect to speaking, two received ED, six earned BD and two, DS.

Overall, in the final term evaluation, control students also progressed.

Concerning the first aspect though, continuous learning, students did not show

progress. Six students received ED, three earned BD and one, DS.

In the second aspect, behavior, six students received ED and four, BD.

In the third aspect, homework frequency, nine students received ED and one BD.

Regarding homework performance, nine students received ED and one BD.

In the participation aspect, nine students received ED and one DS.

Concerning speaking, five students received ED, four got BD and one DS. For better comparison, the results are properly observed in Figure 40.

Figure 41 presents the results of the statistical comparisons concerning the Pre-Kinder 2 control group comparison of middle and final term evaluation of the school. The results were also non-statistic significant, except for the behavior (p=0.049) and speaking (p=0.15) components. This result was coherent with what was observed in the classes and exposed by the teacher during the interviews and informal conversations. This group presented some behavior problems that were dealt with throughout the semester.

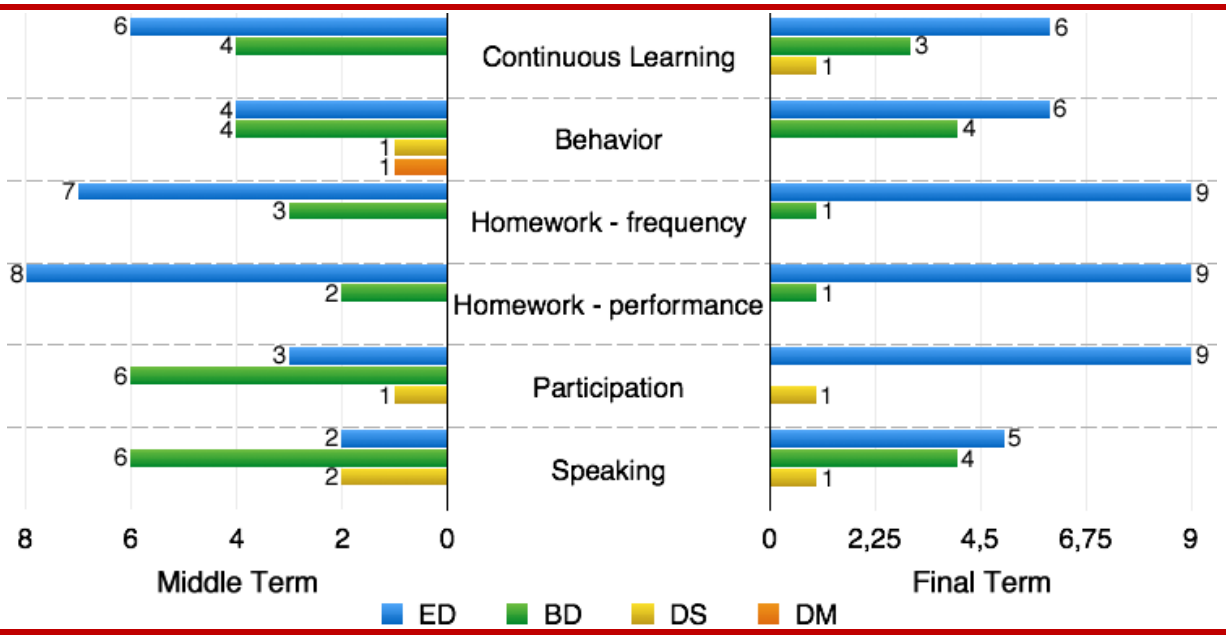


Figure 40: On the left side, Pre-Kinder 2 control group performance in the middle term evaluation. On the right side, Pre-Kinder 2 control group performance in the final term evaluation

7.1.5. Final Test

Since the contents worked with the tool were not predetermined by researchers but crafted with the help of the teachers, only a post-test was elaborated. The teacher worked with contents from 7 units of the book in total. The book consists of nine units. The test applied consisted of 6 questions (19 items, in total), divided in 16 questions about contents worked with the tool and 3 questions related to content not worked with it. The test was applied with both case and control groups.

For this activity, all the students from the case group answered the test (12, in total), while from the control group, nine students did the test (one student was missing).

Figure 42 shows the results of both case and control groups in the post-test. The case group grade is divided in two parts (the

questions whose contents were worked with the ARBlocks and the questions which were not studied with the tool).

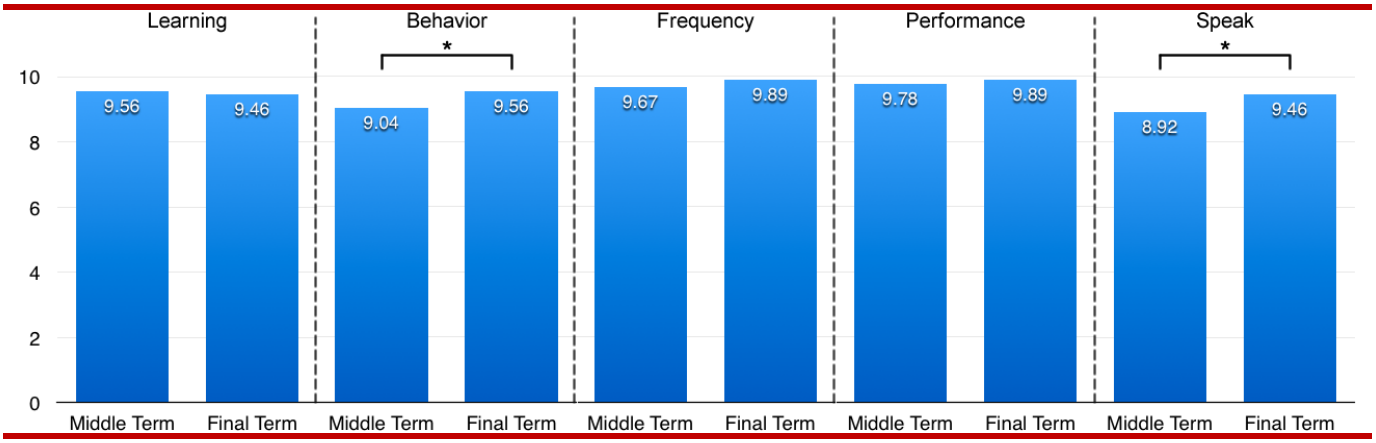


Figure 41 Pre-Kinder 2 control group results for comparison between middle term and final term evaluation. (*p<0.05 based on Paired T- test)

Through the chart, it is possible to see that case group achieved a greater score in the questions using AR. Case group average grade in the AR part was 8.649 while their average grade in the non-AR part was 7.5. The score was even slightly higher than the general score of the control group whose average score was 8.538.

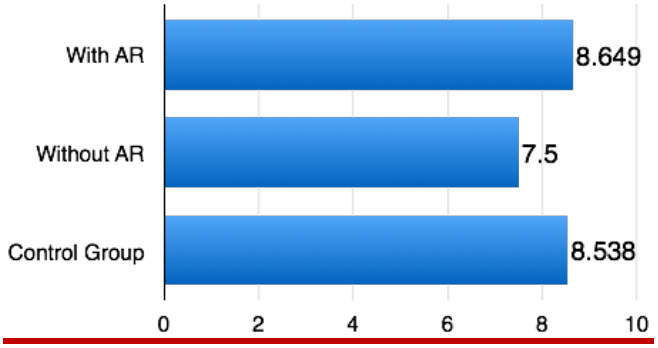


Figure 42: Post-test scores of the Pre-Kinder 2 case and control groups

In this final test applied by the researchers, the Pre-Kinder 2 groups did not present statistic significant differences. In this test, the control group presented slight better scores than the experiment. Although the observations and the other instruments suggested the ARBlocks had a positive impact on students learning, this impact was not statically reflected in this test as shown in Figure 43.

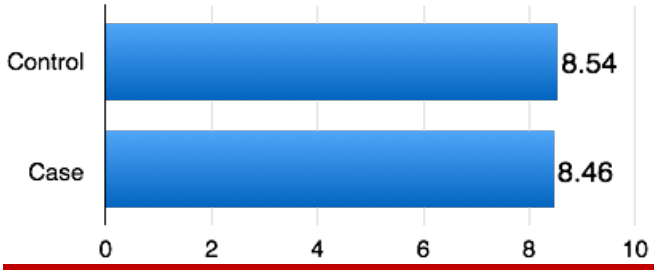


Figure 43: Pre-Kinder 2 group's results for the final test applied by researchers.

7.1.6. Satisfaction Analysis

The satisfaction questionnaire was applied by the researcher who led students to reflect about the experience of using the ARBlocks and emphasized that there was no right or wrong answer but they were free to expose their own feelings.

In the first question that asked if students liked to use the ARBlocks, eleven students colored green and one colored yellow.

Regarding the second question that asked if students liked the activities with the blocks, eleven colored green and one colored red.

Concerning the third statement which was “I learned with the blocks”, eleven colored green and one colored yellow.

The fourth statement led students to reflect if it was easy to use the blocks, eleven students colored green and one colored yellow.

The fifth statement inquired if students would like to play again with the tool. For this aspect, nine colored green, one colored yellow and two red.

The last statement inquired if they would recommend that game to a friend. For this one, eleven answered green and one yellow.

All the answers are detailed in Figure 44. We can see that overall students evaluated the tool positively.

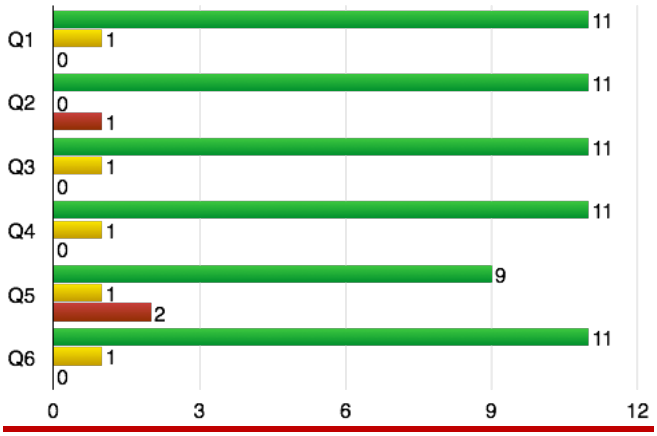


Figure 44: Pre-Kinder 2 answers for the satisfaction questionnaire

7.1.7. Final Interview

In the final interview, the teacher was asked to evaluate the experience of using the ARBlocks in her Pre-Kinder 2 group. The conversation was audio recorded and lasted about ten minutes. She evaluated the experience as positive and constructive for the children since they are already familiar with digital equipment. She argued that the ARBlocks is an extension of that reality. One thing that she mentioned as very positive is the fact that they can touch the blocks and manipulate the information. She claimed that it was excellent concerning the activities.

She mentioned that the ARBlocks perfectly matched with what they were studying. It was easier for her to adapt the activities to the blocks. In addition, she liked the fact that it was possible do add the listening. In her opinion, the tool was simple and easy to use. She argued that she did not need to interfere so much during the activities only to give some guidance regarding what was being requested.

When argued about the positive and negative aspects, she mentioned that one wonderful aspect was the auditory element. She really enjoyed the possibility of combining the visual (writing) to the auditory aspect. She mentioned that students used to move the blocks and wait for the sound. As for negative aspects, she did not remember anything that complicated the process.

Another question was about planning. The teacher considered the process of planning as smooth. She did all the planning basically in two moments. In the beginning of the process, she thought about 6 activities and in the middle of the process, she thought about 6 more.

She explained that she thought about adapting the content seen in class, however, she took into account how was the best way to display the information she wanted to focus on in the blocks.

She also mentioned that she felt great connection among what the ARBlocks offered and what she wanted the students to analyze and to what they are used to work in the classroom. To conclude, she believed planning was a very natural process.

In what concerns children interaction aspects, she thought it was excellent. She claimed that they seemed to be anxious to participate in the activities. She also mentioned one of her students who demonstrated resistance to work with writing in a paper, but felt really motivated to manipulate the ARBlocks. He did not feel intimidated to make mistakes. When the sessions were over, she remembered that some students were still in the classroom playing with the ARBlocks. She also highlighted that students did not have problems even when the projection took longer to appear.

Regarding the dynamic of the activities, she mentioned it was very varied, for instance, she mentioned some examples: activities to make words, pictures to place correctly in a given context and sorting out.

The teacher explained the criteria she used to decide the interaction. First of all, she claimed to think about how to help students in the process of beginning to read and write in English. Second of all, she considered how to best organize the students for the activity, if it was better to make pairs in which the strongest students would help the others or individually (in case the activity was simple for them).

In a nutshell, she evaluated the interaction with the tool as excellent. She mentioned that when there were no activities, the students used to ask if they were going to the room 222 (the activities were done in a different room at the end of every class). According to the teacher, they were anxious to finish the class with the blocks because to them this was like bringing the class to a perfect end. She mentioned that they loved it and when the activities were concluded, they resented that it had finished. In her words, they missed it. In her opinion, the tool contributed to students' learning. She mentioned that she noted that in the student's eyes.

When asked if the tool had any impact in the student's learning, she replied that she noticed learning and that they got used to do the activities and waited for the feedback. This suggests that students were able to work independently.

To conclude, she mentioned that the experience was positive and she resented not having used the tool with her other Pre-Kinder 2 group (the control group). She believed that although they had some behavior issues, they would have benefited from using the ARBlocks. She also highlighted that the projection on the floor is different from the screen and that she supposed that the brightness is more comfortable to students on the floor. She also highlighted that students were able to interfere in the projection through their actions.

7.2. KIDS 1 RESULTS

In this section, it will be presented the results from the Kids 1 groups.

7.2.1. PRELIMINARY INTERVIEW

As with the Pre-Kinder 2 teacher, in the first interview, the Kids 1 teacher was asked about her impressions of the ARBlocks. The interview was audio-recorded and lasted approximately four minutes. She explained that the tool is strongly related to the school's work since technology is already part of their lives. She also believed that it might raise students' attention and, therefore, motivate them.

Regarding how to work with the language abilities, she explained that she saw different possibilities on activities involving matching, sorting out and gap filling. She explained that she had some difficulties in realizing how to work with the reading or speaking skill, for example.

7.2.2. Activities

Due to time constraints, this teacher chose to use the ARBlocks once a week. Therefore, six activities were done with the students. All the activities involved reinforcement of the content seen in class.

The first activity worked with the prepositions ”in” and “on”. Students were presented to some questions (e.g.: where is the bird?) and pictures and had to choose between in and on to complete the sentences.

The second activity was intended to practice furniture vocabulary and the use of the grammar structure “there is/there are”. Students were presented to a picture of a room and a question related to what they could see. They needed to choose the right answer using the blocks to complete a sentence (e.g.: there is a TV in the living room) as seen in Figure 45.

The third activity also involved furniture vocabulary but it was intended to practice the grammar structure of “yes/no” questions. Students saw a question (e.g.: is there a bed in the bedroom?) and answered the question according to the picture associated to it. The answers may be yes, there is or no, there isn’t.

Figure 45: Kids 1 students practicing grammar structure using the ARBlocks

Another activity was related to the structure what are you wearing and to clothing vocabulary. Students were presented to some pictures and asked what the character in the picture was wearing. They chose the correct answer displayed in the blocks.

The last two activities were related to unscramble the names of toys and food vocabulary, respectively. In Figure 46, it is possible to see children practicing food vocabulary.

Figure 46: Kids 1 student using the ARBlocks to unscramble food vocabulary

7.2.3. Research Diary

As with the Pre-Kinder 2 teacher, firstly, the teacher recorded some basic information such as the track, the time, the date and what was the activity done.

One of the questions was related to what were the skills practiced with the tool.

Teacher mentioned reading, listening, recognition, vocabulary and spelling.

In most of the sessions (5, in total), students interacted with the tool individually. In the first session, they interacted in small groups as a competition.

In all the activities, the teacher interacted with the group during the activity.

In the first four sessions, teacher reported that the students liked the activities. For instance, she mentioned that they were “excited to use the blocks” and “eager to participate and to touch the blocks” and that “they liked and helped each other”.

However, in the last two sessions, she reported that students were “a little bored” and that she felt they need more interaction.

When asked if she observed them learning something using the application, her answers were mostly related to reviewing and reinforcement. She also mentioned that they learned spelling and vocabulary.

In the question about possible difficulties students had faced during the use of the ARBlocks, she argued most of the times they did not had difficulties, except in the clothing activity in which she reported that some pieces of clothing were not so clear.

To conclude, she was asked to freely report any comment or situation. This section was left empty most of the time, however, in the last activity, she mentioned that “maybe the kind of interaction was not adequate for them”.

7.2.4. Academic Achievement

7.2.4.1. Case Group

Through the middle term evaluation, it is possible to observe that overall students presented a good development and command of the language.

In the first requirement, continuous learning, six students received ED, two students received BD and one DS.

As regards to behavior, eight students obtained ED and one, BD.

Regarding homework frequency, seven students received ED, one got BD and, finally, one had DS. Students received these same scores for homework performance.

As for participation, seven students received ED and two received BD.

Regarding speaking, four obtained ED, three got BD and two, DS.

In the final term evaluation, it is possible to observe that students progressed in most of the topics.

For instance, in continuous learning, seven students received ED and two got BD.

As for behavior aspect, the students remained with the same grades which were eight students with ED and one with BD.

For homework frequency, six students received ED and three received BD.

For homework performance, six students got ED, two got BD and one received DS.

When it comes to participation, eight students received ED and one received BD.

Finally, in the speaking aspect, six students received ED and three received BD as presented in Figure 47.

In the middle term evaluation comparison of the Kids 1 groups, there was no significant difference for the aspects, except behavior. Therefore, it is noticeable that they were different concerning this aspect ($p = 0.009$). This may be explained by the fact that the control group is composed by seven boys

and one girl. These students were more agitated as was mentioned by the teacher and confirmed by the observations.

Figure 48 exposes the results of the Kids 1 case group concerning the comparison of middle and final term evaluation of the school. The results were also non-statistic significant, except for the speaking component (p=0.036).

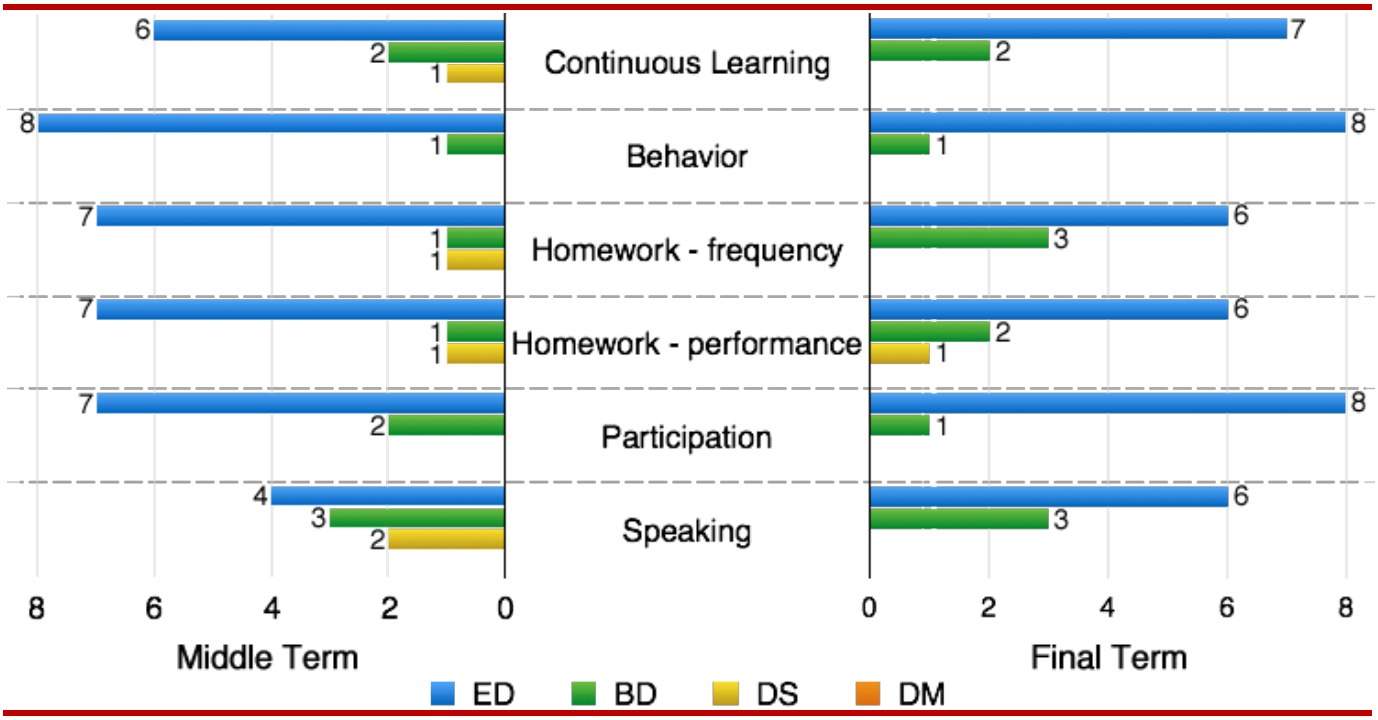


Figure 47: On the left side, Kids 1 case group performance in the middle term evaluation. On the right side, Kids 1 case group performance in the final term evaluation

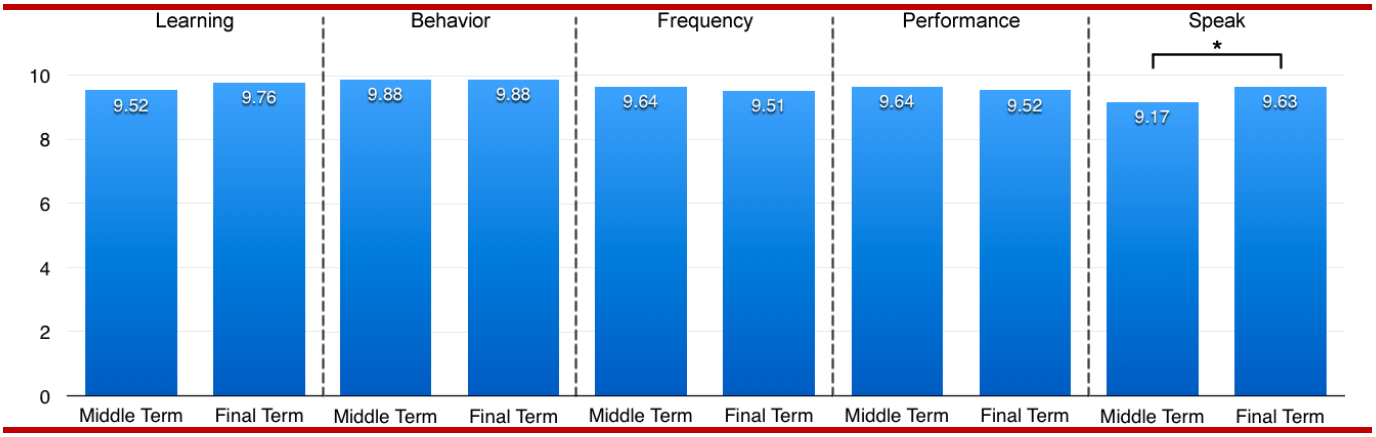


Figure 48: Kids 1 case group results for comparison between middle term and final term evaluation (*p<0.05 based on Paired T-test).

As for the speaking, students from both Kids 1 groups and the Pre-Kinder 2 control group presented statistic significant improvements. This may be explained by the fact that speaking is an ability that generally students have more difficult with since it is difficult for them to practice it outside the school. Hence, teachers constantly reinforce this ability at school. For instance, in the kids 1, it was observed that the teacher had a chart in the class in which she marked every time students used expressions in Portuguese related to contents that they had already learned.

7.2.4.2. Control Group

Through the middle term evaluation, it is possible to observe that overall students presented a good development and command of the language. Although their scores were, in general, lower than the case group.

For continuous learning, two students received ED. Four students got BD and two students received DS.

For behavior, two students obtained ED, five got BD and one received DS.

As for homework frequency, four students received ED, three got BD and one received DS. Those students received the same scores for homework performance.

As regards to participation, seven students received ED and one got BD.

As for speaking seven students got BD and one DS.

Through the final term evaluation, it is possible to see that students overall progressed in almost all categories.

For continuous learning, six students received ED and two remained with DS.

For behavior, four students had ED and four got BD.

In the homework frequency aspect, it is possible to observe that concepts decreased a little. Concerning this aspect, two students received ED, three students got BD and three had DS.

As for homework performance, it is noticeable some improvement. Four students had ED and four got BD.

As for participation, all the students received ED.

Finally, in the speaking category five students received ED and three got BD. All students' scores for both middle and final term evaluation are presented in Figure 49.

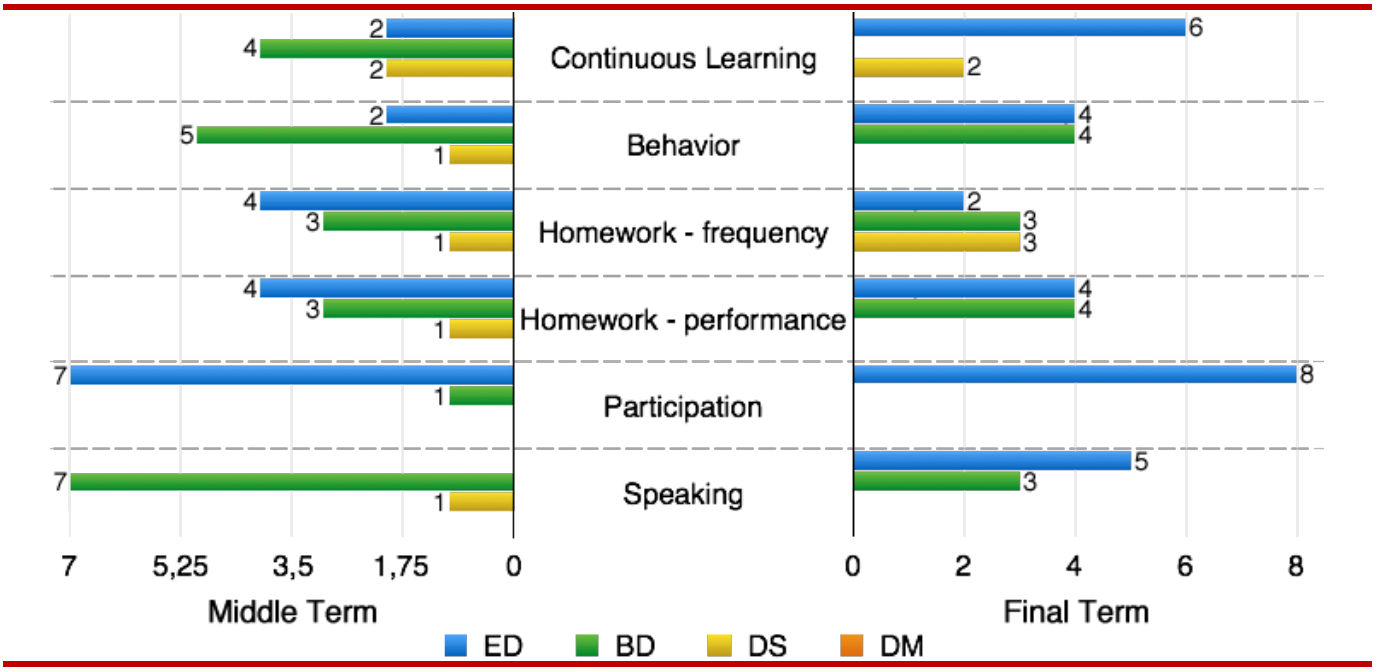


Figure 49: On the left side, Kids 1 control group performance in the middle term evaluation. On the right side, Kids 1 control group performance in the final term evaluation

Figure 50 exposes the results of the Kids 1 case group concerning the comparison of middle and final term evaluation of the school. The results were also non-statistic significant, except for the behavior (p=0.020) and speaking component (p=0.08). The

results concerning behavior were coherent with the other results obtained. As previously exposed, this group had more boys than girls and presented behavior problems. These behavior issues were dealt with throughout the entire semester.

As for the speaking, students from both Kids 1 groups and the Pre-Kinder 2 control group presented statistic significant improvements. This may be explained by the fact that speaking is an ability that generally students have more difficult with since it is challenging for them to practice it outside the school. Hence, teachers constantly reinforce this ability at school. For instance, in the Kids 1, it was observed that the teacher had a chart in the class in which she marked every time students used expressions in Portuguese related to contents that they had already learned.

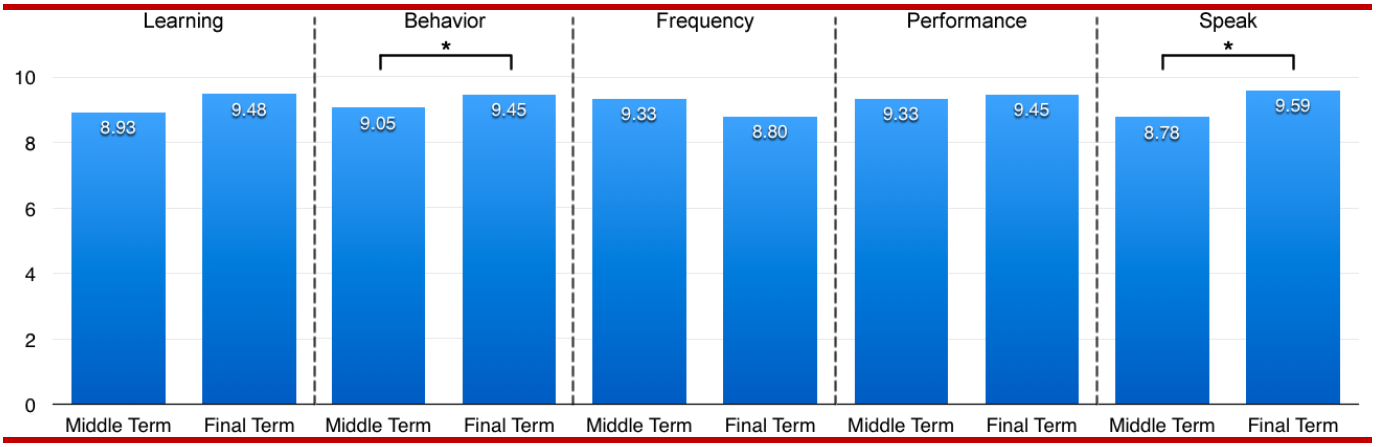


Figure 50: Kids 1 control group results for comparison between middle and final term evaluation (*p<0.05 based on Paired T-test and $^{+}$p<0.05 based on Wilcoxon test).

7.2.5. Final Test

As previously explained, after the work with the ARBlocks, a post-test evaluation was applied. This evaluation consisted of five questions (23 items, in total) related to contents studied with and without the help of the ARBlocks.

The teacher worked with contents from four units of the book in total. The items in the test were divided in 15 items worked with the tool and 8 items related to content not worked with it. The test was applied with both case and control groups. A total of eight students in each class answered the tests.

Figure 51 presents the scores of the post-test in both case and control groups. The case group results are divided between the questions worked with and without AR support. The control group did not need this division since all the questions were worked in the traditional way.

By analyzing the graph, it is possible to see that case group had the same scores in the entire test (8.75). This grade was higher than the grade of the control group (8.152).

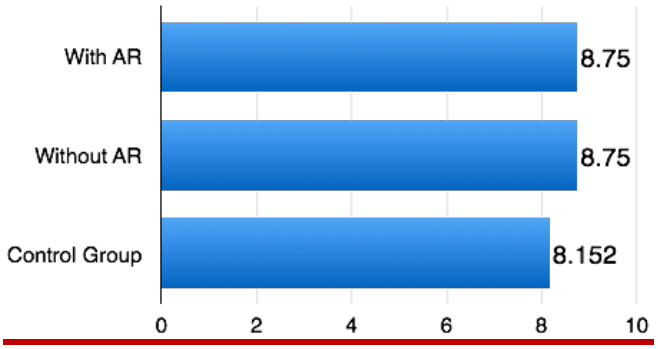


Figure 51: Post-test scores of the Kids 1 case and control groups

As for the final test applied by the researchers in the Kids 1 groups, results presented statistic significant differences as

can be seen in Figure 52. In this case, the experiment group presented better scores (p=0.026). The observations and the other instruments suggested that the ARBlocks had a positive impact on the reinforcement of the content, although the teacher had some reservation regarding the interaction aspects of the tool.

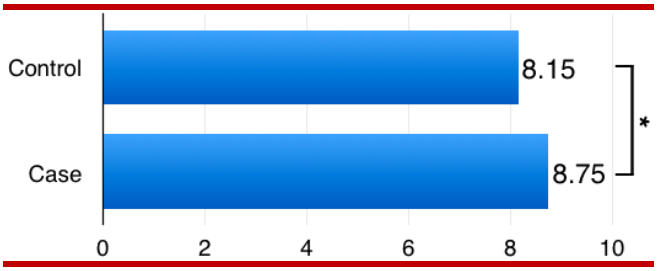


Figure 52: Kids 1 groups' results for the final test applied by researchers (*p<0.05 based on independent samples T-test).

7.2.6. Satisfaction Analysis

As happened with the Pre-Kinder 2 group, the satisfaction questionnaire was applied by the researcher who led students to reflect about the experience of using the ARBlocks. The researcher reminded students that there was no right or wrong answer and that they were free to expose their own feelings. In this group, not all students answered the reflection because some of them were missing. Therefore, we had a total of six responses.

The first question inquired if students liked to use the ARBlocks, five students colored green and one colored yellow.

Regarding the second question which asked if students liked the activities with the blocks, three colored green and three colored yellow.

The third statement was "I learned with the blocks". For this aspect, one student colored green, two yellow and three colored red.

The aim of the fourth statement was to reflect if it was easy to use the blocks, five students colored green and one colored yellow.

Concerning the fifth statement which inquired if students would like to play again with the ARBlocks, three colored green and three colored yellow.

The last statement inquired if they would recommend that game to a friend. For this one, four answered green and two yellow.

The answers for all those statements are detailed in Figure 53.

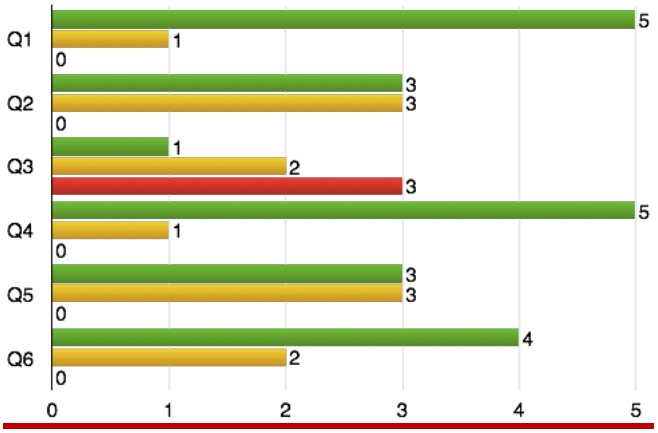


Figure 53: Kids 1 answers for the satisfaction questionnaire

7.2.7. Final Interview

In the final interview, the teacher was asked to evaluate the experience of using the ARBlocks in her classroom. The interview was audio recorded and lasted approximately eleven minutes. She explained that the experience was positive but there is a need to take into account the interaction aspect with the tool which could have been done differently to their reality of an English language school. She mentioned that all the tools involving technology are positive and she does not oppose to use them in the classroom.

When asked about the positive and negative points of using the ARBlocks, she

mentioned that the main issue was the interaction. She believes it was repetitive. The activities ended up being just about vocabulary recognition. She stated that the technology is interesting but it needs to be expanded to explore more reading or speaking, for instance. She believes it is a valid technology. It is something to add in the classroom. However, she believes it should be seen in a different way since students are used to competition.

Regarding planning, teacher claimed she did not know that she needed to plan the lessons and include the ARBlocks. She thought she needed to explain her lesson plan to programmers and they would adapt the activities to be used in the ARBlocks. Therefore, she thought about her lesson plan and how some activities could be done with the blocks. She mentioned that it would have been different if she had thought the opposite way (to think about an activity for the ARBlocks).

The teacher explained that the problem was not only the interaction but group work in which students would interact at the same time. She mentioned that during the activities, students who were not working were bored even though she tried to ask them to help and join the group.

Regarding the impact of the tool, she explained that she would not actually call it impact, but reinforcement. She claimed that the tool was a reinforcement of content previously seen in class. She argued that the tool came to complement, to help, as a novelty to help in the classroom.

The teacher mentioned that the tool had its contribution, that it was positive. Nevertheless, she lists some points of concern, such as the size of the projection on the floor (she believed it was small) and the limitation of the blocks.

As for contributions, she mentioned the reinforcement of language structure and the combination of the visual aid and the ability to touch and test hypotheses.

To conclude, she mentioned that in order to bring this tool to her school, it would be necessary to rethink some questions, such as the projection and the interaction. She mentioned that, in general, she really liked the tool and was curious to know how different people would work with that in her reality.

7.3. DISCUSSION

As regards to the results of the evaluation, we can divide them in two parts, the Pre-Kinder 2 and Kids 1 results. The ARBlocks was used in both groups as a supplemental tool in the classroom to foster student's learning.

In the Pre-Kinder 2 groups, although it was not possible to work specifically with writing, the results supported the hypothesis that the ARBlocks can help to motivate students and foster the development of their language skills.

In the post-test evaluation (applied after using the tool), the case group achieved a greater score in the questions using AR. Case group average grade in the AR part was 8.649 while their average grade in the non-AR part was 7.5. The score of the AR part of the test was slightly higher than the control group (8.538). Statistical comparisons showed that the results were non-statistic significant. In this test, the control group presented slight better scores than the case group. Although the observations and the

other instruments suggested the ARBlocks had a positive impact on students learning, this impact was not statistically reflected in this test.

As concerns student's academic achievement, through their regular evaluation, it was possible to observe that students from both groups (case and control) had similar language levels.

Although there were slight differences. For instance, in the continuous learning aspects, the control group presented slightly better scores (six students with ED and four with BD) while the scores for the case group were three students with ED and nine with BD.

As the teacher previously mentioned the control group presented some behavior problems, therefore, in this aspect, case group presented higher scores (four students with ED, four with BD and four with DS) while for the control group, four students with ED, four with BD, one with DS and one DM.

In the homework frequency aspect, case group presented better scores (ten students with ED and two with BD). As for the control group, we have seven students with ED and three with BD.

However, when it comes to homework performance, we have the opposite situation. Control group students presented better scores (eight students with ED and two with BD) while for the case group, four students achieved ED and eight, BD. As for participation, case group presented better scores (four presented ED and eight, BD) while in the control group, three achieved ED, six, BD and one, DS.

As for speaking skills, control group presented a slightly better score with two students with ED, six with BD and two with DS. As for the case group, we had three students with ED, four with BD and five with DS.

In the final term evaluation, it was noticeable that students overall progressed in their learning scores.

In the continuous learning aspect, case groups achieved satisfactory progress in the scores. Five students achieved ED and seven, BD.

On the control group, students remained with the same scores, expect from one student who regressed one score, so the results were six students with ED, three with BD and one with DS.

For the behavior aspect, the case group presented nine students with ED, one with BD and two with DS. The control group, on the other hand, presented six students achieved ED and four, BD.

In the homework frequency aspect, the results for case group were: eleven achieved ED and one, BD. In this aspect, the results of the control group were: nine students with ED and one with BD. The same scores were achieved in homework performance in both groups.

As regards to participation, case group presented the following scores: ten students with ED and two with BD while the control group presented nine with ED and one with DS.

Finally, in the speaking aspect, case group presented three students with ED and nine with BD while for control group the results were five students with ED, four with BD and one with DS.

As could be noted, the case group achieved better scores in four aspects, behavior, homework frequency, homework performance, participation and speaking.

Statistic comparisons between the middle and final term evaluation showed that most of the categories were not significantly different. For the case group, only the homework-performance presented statistical difference (p=0.025). This result may suggest that the use of the tool had a positive impact on student's performance regarding the topics studied. The other research instruments used reinforced this positive impact.

For the control group, the results were also non-statistic significant, except for the behavior (p=0.049) and speaking (p=0.015). This result was coherent with what was observed in the classes and exposed by the teacher. This group presented behavior problems that were dealt with throughout the semester.

Regarding the interaction aspects, the teacher preferred to work with students individually (seven times), although she also worked with the students in pairs (four times).

Through the interview, the teacher highlighted that she really liked the fact that the tool allows students to manipulate the content (interfere in the projection) and its association with the sound. This is aligned with Vygotsky's theory since he highlights that children's understanding comes not just from information, but also from what they see and hear especially they have a chance to interact with (MOONEY, 2013).

Additionally, studies have shown that young children are less likely to focus on the vocabulary or pronunciation errors of others or to correct them when learning a foreign language. Hence, the auditive feedback may help in this aspect.

According to the teacher feedback, the tool helped students to learn different contents and abilities. For instance, in one of the sessions, the teacher reported students learning a social ability, that is, "they learned how to wait their turn". This is aligned with Dewey's theory since this author emphasizes the need to teach children how to live socially (MOONEY, 2013). This aspect was also seen in the pilot study when students had to collaborate in small groups in order to fill in a nursery rhyme.

The ludic aspect added by the tool seemed to be important to reduce anxiety and student's resistance. As the teacher reported that one of her students seemed to be anxious to participate in reading activities in the classroom but felt really motivated to work with the ARBlocks. She reported that he did not feel intimidated to make mistakes.

The children's opinions through the questionnaire reinforced this conclusion since almost all children colored green for the questions regarding satisfaction issues.

It was observed that the tool helped to enhance motivation among students, the teacher mentioned one student that always wanted to stay in the classroom finishing activities even when the time was out. This aspect was also observed through video recording and observations when some students always wanted to take part in the activity and the teacher needed to call their attention to wait for their turn. The teacher also mentioned that they were anxious to go to the room where the activities took place, in her opinion, to finish the class with the AR activity was like bringing it to a perfect end for them.

As regards to Kids 1 groups, we can say that the results of introducing the ARBlocks was not as good as expected.

In the post-test evaluation (applied after using the tool), case group achieved the same score in the two parts of the texts that corresponded to questions worked with and without the tool (8.75). Their general score was higher than the one achieved by the control group (8.152).

Statistic comparisons revealed that these results were statistically significant. The case group presented better scores (p=0.026). The observations and the other instruments suggested that the ARBlocks had a positive impact on the reinforcement of the content, although the teacher had some reservation regarding the interaction aspects of the tool.

Although the teacher mentioned that students in the case group were starting in the program and only one student was repeating the level, through student's academic achievement, it was possible to observe that, in general, case group students presented better scores than the control group.

In the continuous learning aspect, case group students had six students with ED, two with BD and one with DS. On the other hand, control group presented two students with ED, four with BD and two receiving DS.

For behavior, case students presented eight students with ED and one, BD as opposed to two students that obtained ED, five who got BD and one that received DS in the control group.

Regarding homework frequency, in the case group, seven students received ED, one got BD and, finally, one had DS. While in the control group, four students received ED, three got BD and one received DS. The scores were the same in both groups regarding homework performance.

As for participation, the results were equivalent, since in case group, seven students received ED and two received BD while seven students received ED and one got BD in the control group.

Regarding speaking, four obtained ED, three got BD and two, DS in the case group while, in the control group, seven students got BD and one DS.

Overall, both case and control groups presented some progress in almost all categories. Nevertheless, both groups presented decrease in the homework frequency. As for the case group, there was also a slight decrease in homework performance as well. The scores for case group in the middle term for this aspect were seven students with ED and one with BD. While in the final term evaluation, six students got ED, two got BD and one received DS for the same aspect.

Statistical comparisons of academic achievement of the case group showed that the results were non-statistic significant, except for speaking (p=0.036). For the control group, results were also non-statistic significant, except for behavior (p=0.020) and speaking (p=0.08). The results concerning behavior were coherent with the other results obtained. This group presented behavior issues, which were dealt with throughout the group presented statistic significant improvements. This may be explained by the fact that speaking is an ability that is generally more difficult for students to practice outside the school. Hence, teachers constantly reinforce this ability at school. For instance, in the Kids 1, it was observed that the teacher had a chart in the class in which

she marked every time students used Portuguese expressions.

As for the interaction aspect, teacher 2 indicated that although the tool supported only a limited interaction, the combination of the visual aid and the ability to test hypothesis was a positive aspect. The teacher mentioned that the contribution was related to reinforcing the contents seen in class.

The students of this class were really used to competition games which were not worked with the tool. She tried a competition game in the first activity. However, in the other sessions, the activities proposed were individual. During the interview, she also mentioned the difficulty to work in groups using the tool.

These results were also corroborated by students through the satisfaction questionnaire since we had a high number of students answering yellow or even red for some of the statements.

Since we aim to discuss the evaluation process and intend to generate some guidelines for future evaluations, it is important to discuss the evaluation process itself and the factors that might have influenced it.

It is important to highlight that some characteristics of the ARBlocks contributed to its acceptance by teachers and students, such as manipulation of the content and its association with the sound. Indeed, this is an important aspect for young children learning as shown in the review of the literature. Different authors highlight that children learn better through sensory experiences. This aspect was emphasized as positive by the teachers during the interviews.

The ludic aspect is also an important aspect as pointed out by the teachers. According to Piaget's theory, as previously exposed, play is an important part of children's learning.

One important aspect of our evaluation was the process of adaptation of the new technology into the learning environment. Regarding this aspect, Pre-Kinder 2 teacher argued that it was easy to adapt content for the tool. She did it basically in two different moments. She mentioned that her first thought was always on how to contribute to student's development and, then, how to better arrange them for the activity.

She mentioned that she did not need to interfere a lot during the activity which may have freed her to better observe students. This aspect is aligned with Montessori's ideas that teachers need to teach little and observe much. Through observation, it was noticeable that she helped them most of the time to understand the writing commands of the activities. This was natural since they were still learning how to read and recognize words.

On the other hand, as was made clear during the interview, we experienced a communication problem with the other teacher since she stated that she was not aware that she needed to plan for the activities ahead of time. Therefore, she claimed to have prepared the activities for her lessons and, then, adapted them for the tool. She acknowledged that this may had influenced the results.

Regarding the interaction, it was noticed that one teacher enjoyed the interactive features allowed by the tool while the other demonstrated to have problems with them

since she had difficulties in elaborating group activities with the tool.

This is an important feature to consider since when using technology devices in the classroom it is important for teachers to have all the students working otherwise they might lose student's attention. The Kids 1 teacher also emphasized the need to do competition activities and that she did not feel it was possible to do it with the tool.

This discussion brings interesting insight about the importance of considering teacher's routines and teaching styles in the use of different tools. Since the other teacher did not seem to miss those aspects as much as she did.

One aspect that is important to emphasize is that the degree of engagement of teachers may have huge impact on the results of research involving the use of technology in schools.

Hence, it is imperative to have high engagement of teachers which is not always possible since it demands efforts from them and may be time consuming.

As regards to our research, one of the teachers was highly engaged while the other one due to time constraints was moderately engaged. Therefore, in these groups the tool was applied only once a week instead of twice a week.

Therefore, the impact may be limited in the Kids 1 group since they used the tool for less time when compared to the Pre-Kinder 2 groups. Although results suggested that they needed more varied activities and interaction to feel motivated to use the AR tool.

As regards to the instruments used, we can say that the use of the research diary helped to capture teachers' memories right after the experience which helped researchers to better understand the process and avoid memory decay. Particularly with the Kids 1 diary, it was interesting to see that students found the tool attractive and motivating in the first four sessions and in the last two started to get a little bored, aspect that was properly discussed during the interview with the teacher. These positive reactions were not appropriately captured during the interview.

This instrument also allowed to capture particular concepts that were reinforced by the students according to the teacher's observations.

The use of the post-test was also important to have a better understand of the impact of the tool in students' development. Moreover, the use of student's academic achievement was very useful since it provided an overview of student's previous knowledge as well as a better overview of the entire process covered along the semester. Therefore, it consisted as a formative evaluation rather than just a punctual measurement.

Finally, the satisfaction questionnaire used was based on a similar questionnaire taken by students in the school. Therefore, students are used with the model. Nevertheless, it is arguable that students that young, in general, may not be mature enough to answer such a type of questionnaire. In addition, this type of instrument also has its limitations. For instance, it is not possible to further understand the reasons behind the answers.

As for our research, we attempted to mitigate those flaws by using researchers' observations as well as teachers opinions.

Although results must be taken lightly since the ARBlocks was not used to work specifically with all the abilities involved in the evaluation, results suggest that the use of ARBlocks may have contributed to student's overall development of both groups although the impact seemed to be higher in the Pre-Kinder 2 than in the Kids 1 group. The use of the multiple instruments both quantitative and qualitative helped the researchers to have a better overview of possible reasons for these results and, therefore, better understand the potential and limitations of the tool analyzed.

8. CONCLUSION

Figure 54: Pre-Kinder 2 student using to ARBlocks to unscramble a color name

The present work intended to provide some reflections about the issues involved in the process of evaluating AR educational technology.

A systematic review was carried out in order to understand how researchers are addressing to this question. Through this work, we could identify AR's potential to be applied in classrooms.

Developments in this technology have enabled researchers to develop more tools in the field of education and to evaluate them. AR has unique affordances that can impact the learning experience.

It was observed that most of the studies investigated not only the learning aspects but usability properties as well.

As technology matures, researchers are increasingly concerned with how to properly evaluate them. Some studies took into account local curriculum or needs pointed out by experts in order to develop proper AR educational tools. One of the studies took into account teacher's experiences.

Multiple metrics were also used by some researchers in order to better understand the complexity of the learning scenario. We found out that only few works applied multiple metrics when it comes to evaluate learning effects. Most of them used multiple metrics, but in order to evaluate different aspects than learning, such as usability and learner's satisfaction.

After the lessons learned through the review, we also evaluated an AR tool especially developed to children education, the ARBlocks in two different contexts.

A pilot study and a final semi-experimental study was conducted in two different institutions, a public and a private one in order to investigate the impact of AR technology in the educational field.

We designed a quasi-experiment in order to examine the impact of the ARBlocks in the Language teaching field (Portuguese and English).

This evaluation encompasses multiple metrics, both qualitative and quantitative as well as the involvement of the teachers in elaborating suitable activities for their classroom needs.

Results from the pilot study supported the hypothesis that the ARBlocks can help students to learn. It was noticed that in a smaller period of time, the children reached a good progress concerning their writing development.

It was interesting to notice that the teacher divided students into stations to work with the tools and with the other activities she had prepared. Additionally, student's own productions were used in some of the activities, such as drawings and audio files.

The multiple metrics used, such as the teacher's regular evaluation, the post-test applied after the use of the tool and the interviews supported that result.

It was noticed that the teacher was very enthusiastic with the use of the tool and she believed that the students were extremely engaged in the activities. It was also noticeable that students got attached to the tool and named it "the robot".

As for the final study using the ARBlocks, we had two sets of results.

In the Pre-Kinder 2 groups, the results also supported the hypothesis that the ARBlocks can help to motivate students and foster the development of their language skills.

In the post-test evaluation (applied after using the tool), case group achieved a higher score in the questions worked with the tool. Their general score was slightly superior to the one achieved by the control group. Although the observations and the other instruments suggested the ARBlocks had a positive impact on students learning, this impact was not statistically reflected in this test since control group presented slight better scores than the experiment.

Nevertheless, a positive impact of the tool was supported by the other metrics used, such as: student's academic achievement, observations, interviews, research diary and student's satisfaction questionnaire.

The teacher mentioned that the tool helped students to learn different contents and abilities. She highlighted the ludic features provided by the tool as she reported that one of her students seemed to be anxious to participate in reading activities in the classroom but felt really motivated to work with the ARBlocks. She reported that he did not feel intimidated to make mistakes. Therefore, she provided a positive feedback regarding the tool and its use.

On the other hand, for the Kids 1 group, the results of introducing the ARBlocks was not as good as expected.

In the post-test evaluation (applied after using the tool), case group achieved the same score in the two parts of the texts that corresponded to questions worked with and without the tool. Their general score was higher than the one achieved by the control group. In this case, the case group presented better scores ($p=0.026$).

The observations and the other instruments suggested that the ARBlocks had a positive impact on the reinforcement of the content, although the teacher had some

reservation regarding the interaction aspects of the tool.

In fact, the students of this class were really used to competition games which were not worked with the tool. She tried a competition game in the first activity. However, in the other sessions, the activities proposed were individual. She reported to have difficulty to work with the tool doing competition games and group work.

These results were also corroborated by students through the satisfaction questionnaire since we had a high number of students answering yellow or even red for some of the statements.

Overall, it was noticeable that, as pointed out by other researchers, student's benefited from the interaction between real and virtual environments as well as from the use of a tangible interface metaphor, which is a valuable aspect for young children who need to manipulate and explore concrete material (MOONEY, 2013). All the teachers involved in the study pointed out as positive the possibility for children to manipulate the content.

It was also evident that the ARBlocks helped to stimulate student's prior knowledge as was evidenced that all teachers used the tool to reinforce content already seen in class. Thus, the ARBlocks provided an additional practice for students. The tool was not used to introduce a new content with them. Results also showed that it may helped to catch student's attention in a fun and engaging way. Pilot teacher and Pre-Kinder 2 teachers exemplified this when they reported of some students who felt at ease during the sessions and, therefore, more prone to practice the reading skill. In addition to that, teachers highlighted the importance of the feedback student's received during the exercises which fostered their independence.

It was noticeable that the tool helped in the beginning stages of the reading skill such as word-recognition, structural knowledge (understanding language structure) and vocabulary recognition. Teachers explored different types of activities, such as: gap-filling, matching, sorting and forming words.

Another feature that was positive was its flexibility since it enable teachers to create different activities, even using student's own productions such as in the activities proposed by the pilot teacher which used student's own drawings, pictures and recordings. Besides, teachers also mentioned as positive the combination of different language skills, such as the association of the reading and the listening. As for a limitation, it was not possible to work with writing skill and reading was possible for small texts such as nursery rhymes, which is appropriate for young learners beginning their literacy process.

As concerns to collaboration, our results suggest that interaction issues can be improved in the tool in order to support more interaction among students through group work or competition activities. This is an important issue since teachers may have to manage large classes on their own and, therefore, they need to keep all students busy during the activities. As for challenges raised, it is important to develop a hardware set that can be easily adaptable to schools and could be applied in the same room as teachers. During our experiments, the sessions were in a different room, which caused students to move from one room to another, the pilot teacher described this process as both positive and negative, since

students may be disturbed by the moving process and it can be difficult for them to get them to focus. A more portable hardware is necessary in order for teachers to easily set them in their classrooms.

As one may expect, AR systems may fail. This happened some times, but, it did not interfere in the development of the activities.

As for the activities preparations, it was noticed that although varying in levels of engagement due to time constraints, the teachers participating in the study were open to new technology and eager to have different activities in their classrooms. An authoring tool is an essential step in the ARBlocks application since it will enable teachers to craft their own activities. Additionally, it is important to provide some sharing among them since teachers may have difficulties in visualizing new activities for their students. This is important, hence, teachers are ultimately responsible for conducting student's learning and use of the tool. In the study, it was observed that student's opinion about the tool was somehow connected to the teacher's opinion and use of it.

As for the evaluation process, we argued for the benefits of involving teachers in crafting activities directly related and adapted for the student's needs. This is a challenging type of evaluation since teachers are not merely subjects but actual partners in the process. This can also be a problematic aspect because it requires a different level of organization and involvement from the teachers who are sometimes overwhelmed with their everyday errands.

Although challenging, we believe this involvement may bring important insight for both researchers and developers who need to understand the users' needs and challenges in order to develop perfectly adaptable solutions for them. We believe that it is important to consider a range of aspects involved in the use of an AR tool in the classroom. For instance, it is important to consider the classroom routines as well as how to involve teachers.

Its is also important for tools to be flexible to adapt to different realities and enable different sorts of interaction, varying from less-controlled to more-controlled activities as well as collaboration among peers.

Another important aspect is that the tool needs to clearly deliver value for the teachers. Teachers need to see utility in the incorporation of a given technology. Hence, this may enhance their engagement in the process.

Regarding the use of teacher's regular evaluation, we believe that it is a good way to have an overview of the impact of the tool in the student's development during the semester or school year.

The use of multiple metrics both qualitative and quantitative is also favored since it allowed researchers to have a better overview of the big picture and counteract for weakness of individual methods as well as better understand problems that might have occurred.

For instance, the use of the research diary helped to capture teachers' memories right after the experience which helped us to better understand the process and avoid memory decay.

As regards to the ARBlocks, we received relevant technical insight and feedback for the tool. Tracking techniques were improved so the ARBlocks could properly work on less

controlled environments such as the schools where the application was used.

In addition, the process of creating new applications was modularized. The modularization enables the reuse of parts of an activity in the creation of a different one which eases the task of creating new applications with the ARBlocks. This modularization is the first step in the development of an authoring tool.

Finally, a set of applications was created which enables the creation of several educational activities by changing only the content. Currently, the library has six applications.

Finally, the results provided by the Kids 1 teacher suggests that it is important to design tools that are flexible enough to provide varied activities and types of interaction for students and teachers.

We concluded that the tool was useful to foster a range of language abilities, although, it offered some limitations for others. AR, in this case, specifically, tangible AR, may help to engage students and involve their whole body in the learning process, thus, encouraging kinaesthetic learning. According to researchers, this is an important characteristic of AR tool along with the possibility to see content in a 3D perspective.

One positive aspect was the fact that many teachers as well as coordinators in the schools were interested in the new technology and during classes' intervals, researchers were able to show the technology to different teachers and discuss the possibilities of using AR tools in different group levels. Most teachers have never heard about AR technology and were willing to understand and learn it.

We are aware that this may not be the reality in all schools but it was a promising indicator of AR tool's potential.

As for limitations, one problem with our study is that it was not possible to have a larger sample of students of the same age or studying the same content which would have reinforced our conclusions. Authors are also aware that the age of Kids 1 (who were a little older) may have interfered in the results. Therefore, results should be taken lightly since individual samples of the groups were restricted and most statistic results were not significant.

Nevertheless, we provided some insight on aspects regarding the evaluation of AR technologies and particularly on the use of the ARBlocks in different contexts and situations.

The use of multiple metrics both quantitative and qualitative for evaluating educational aspects as well as the involvement of the teachers in the entire evaluation process are considered preliminary guidelines for evaluating AR technology with young students. It was observed that the way teachers used a given technology might have a huge impact on student's perception and outcomes regarding its use. As concerns this work, this evaluation provided rich insight regarding the potential benefits as well as limitations of this AR tool.

Thus, evaluations considering these aspects might be expanded to different AR tools as well as different contexts, which may provide data for a deep discussion regarding this topic.

It is also important to admit that using the evaluation of the teacher in the process might be problematic since the standards

may vary from school to school. In addition, it was not possible to measure the exact influence of the tool in each aspect of the scores result.

It could also be interesting to compare students' academic evolution in the period using the tool with students from previous years taking the same course.

With the ARBlocks, it was also possible to work not only with content abilities but with behavior aspects as well, for instance, in the research diary, when asked if she observed students learning something, the Pre-Kinder 2 teacher reported that "they learned especially how to wait for their turn". Future works might investigate how AR tools could influence students' competencies and social abilities.

In addition, as a future work, we expect to develop the authorship tool for the ARBlocks since this is an important aspect for teachers to create their applications independently.

9. REFERENCE

ADAMS, M. J.; FOORMAN, B. **Consciência Fonológica em Crianças Pequenas**. Porto Alegre: Artmed, 2005.

ALMGREN, J. et al. **Tangible User Interface for Chemistry Education:** Visualization, Portability and Database. SIGRAD 2005. [S.l.]: [s.n.]. 2005. p. 19-24.

AMOR, B. Experience Art through Relative Realities with Samsung Digital Gallery at Yuchengco Museum. **AdoboTech – Experience Art through Relative Realities with Samsung Digital Gallery at Yuchengco Museum**, 11 August 2014. Disponivel em: <http://www.adobotech.net/experience-art-through-relative-realities-with-samsung-digital-gallery-at-yuchengco-museum/>. Acesso em: 29 June 2015.

AZUMA, R. et al. Recent Advances in Augmented Reality. **IEEE Computer Graphics and Applications**, November/December 2001.

AZUMA, R. T. A survey of Augmented Reality. **Presence: Teleoperators and Virtual Environments** , v. 6, n. 4, p. 355 - 385, August 1997.

BAKER, F. W. Media Literacy: 21st Century Literacy Skills. In: JACOBS, H. H. **Curriculum 21:** essential education for a changing world. Alexandria: ASCD Member Book, 2010. p. 133-152.

BALOG, A.; PRIBEANU, C. The role of perceived enjoyment in the students acceptance of an augmented reality platform: a structural equation modelling approach. **Studies in Informatics and Control**, Bucharest, v. 19, n. 3, p. 319-330, September 2010.

BILLINGHURST, M. Augmented Reality in Education. **New Horizons for Learning**, Seattle, 2002.

BILLINGHURST, M.; DUNSER, A. Augmented reality in the classroom. **Computer**, v. 45, n. 7, p. 56-63, July 2012.

BILLINGHURST, M.; GRASSET, R.; LOOSER, J. **Designing Augmented Reality Interfaces**. ACM SIGGRAPH Computer Graphics. Los Angeles: ACM. February 2005. p. 17-22.

BILLINGHURST, M.; KATO, H. Collaborative Augmented Reality. **Communications of the ACM - How the virtual inspires the real**, v. 45, n. 7, p. 64-70, July 2002.

BILLINGHURST, M.; KATO, H.; POUPYREV, I. The MagicBook— Moving Seamlessly between Reality and Virtuality. **Computer Graphics and Applications**, v. 21, n. 3, p. 2-4, 2001.

BIMBER, O.; RASKAR, R. **Modern Approaches to Augmented Reality**. ACM Siggraph 2006. Boston: ACM. 2006.

BRAGANÇA, A.; CARPANEDA, I. **Porta Aberta - Letramento e Alfabetização 1º ano**. [S.l.]: FTD, 2008.

BROLL, W. et al. Toward next-gen mobile AR games. **Computer Graphics and Applications**, v. 28, n. 4, p. 40–48, 2008.

BROWN, H. D. **Principles of Language Learning and Teaching**. New York: Pearson Education Company, 2000.

BROWN, H. D. **Teaching by Principles:** An Interactive Approach to Language Learning. [S.l.]: Pearson- Longman, 2007.

CAI, S.; WANG, X.; CHIANG, F.-K. A case study of Augmented Reality simulation system application in a chemistry course. **Computers in Human Behavior**, v. 37, p. 31-40, 2014.

CAI, S.; WANG, X.; CHIANG, F.-K. A case study of Augmented Reality simulation system application in a chemistry course. **Computers in Human Behavior**, v. 37, p. 31–40, 2014.

CAMERON, L. **Teaching Languages to Young Learners**. [S.l.]: Cambridge University Press, 2001.

CARMIGNIANI, J. et al. Augmented reality technologies, systems and applications. **Multimed Tools Appl**, n. 51, 2011. 341–377.

CLARK, A.; DÜNSER, A.; GRASSET, R. **An Interactive Augmented Reality Coloring Book**. SIGGRAPH Asia 2011. Hong Kong: ACM. 2011.

COHEN, L.; MANION, L. **Research Methods in Education**. Third. ed. London: Routledge London and New York, 1989.

CONSORTIUM, T. N. M. **Horizon report Technical Report**. Horizon report. California. 2011.

CROMPTON, P. Evaluation: A practical guide to methods. In: STONER, G. **LTDI Implementing Learning Technology**. Edinburg: Institute for Computer Based Learning, 1996. p. 66-73.

DEMARRAIS, K. Qualitative interview studies: Learning through experience. In: DEMARRAIS, K.; LAPAN, S. D. **Foundations for research**. Mahwah: Erlbaum, 2004. p. 51-68.

DEXTER, S. eTIPS - Educational Technology Integration and Implementation Principles. In: ROGERS, P. L. **Designing Instruction for Technology-Enhanced Learning**. [S.l.]: IGI Global, 2002. p. 56-70.

DI SERIO, Á.; MARÍA IBÁÑEZ, B.; KLOOS, C. D. Impact of an augmented reality system on students' motivation for a visual art course. **Computers & Education**, v. 68, p. 586-596, 2013.

DRUIN, A. The Role of Children in the Design of New Technology. **Behaviour and Information Technology**, v. 21, p. 1-25, 2002.

DÜNSER, A. et al. **Applying HCI principles to AR systems design**. 2nd International Workshop on Mixed Reality User Interfaces: Specification, Authoring, Adaptation. Charlotte: [s.n.]. 2007.

DÜNSER, A.; BILLINGHURST, M. Evaluating Augmented Reality Systems. In: FURHT, B. **The Human Interface Technology Laboratory**. New York: Springer, 2011. p. 289-307.

DÜNSER, A.; GRASSET, R.; BILLINGHURST, M. **A survey of evaluation techniques used in augmented reality studies**. ACM SIGGRAPH ASIA 2008. Singapore: ACM. 2008.

EASTERBROOK, S. et al. Selecting Empirical Methods for Software Engineering Research. **Guide to Advanced Empirical Software Engineering**, p. 285-311, 2008.

EDIGER, A. Teaching Children Literacy Skills in a Second Language. In: CELCE-MURCIA, M. **Teaching English as a Second or Foreign Language**. Boston: Heinle Cengage Learning, 2001.

ERIKSON, E. **Childhood and Society**. New York: W. W. Norton & Company, (1950)1963.

FERREIRO, E. **Reflexões sobre alfabetização.** São Paulo: Cortez Editora, 1985.

FITZPATRICK, A. **Analytical survey information and communication technologies in the teaching and learning of foreign languages: State-of-the-art, needs and perspectives. unesco institute for information technologies in education**. UNESCO Institute for Information Technology in Education. Moscow, p. 106. 2004.

FOXLIN, E. et al. **FlightTracker:** A Novel Optical/Inertial Tracker for Cockpit Enhanced Vision. Third IEEE and ACM International Symposium on Mixed and Augmented Reality. Arlington: IEEE Computer Society. 2004. p. 212 - 221.

FURIÓ, D. et al. Evaluation of learning outcomes using an educational iPhone game vs. traditional game. **Computers & Education**, v. 64, p. 1–23, 2013.

GOOGLE. Google Glass, 2015. Disponivel em: <http://www.google.com/glass/start/>. Acesso em: 29 May 2015.

GOOGLE. Google Glass. Disponivel em: <https://www.google.com/glass/start/>. Acesso em: 18 December 2014.

GRABE, W. Current Developments in Second Language Reading Research. **TESOL Quaterly**, v. 25, n. 3, p. 375-406, 1991.

HARMER, J. **The Practice of English Language Teaching**. London: Longman, 2001.

HIT LAB NZ. HITLabNZ - CityViewAR. Disponivel em: <http://www.hitlabnz.org/index.php/products/cityviewar>. Acesso em: 2014.

HUA, H. et al. **Using a headmounted projective display in interactive augmented environments**. IEEE and ACM International Symposium in Augmented Reality. New York: [s.n.]. 2001. p. 217-223.

JACELON, C. S.; IMPERIO, K. Participant Diaries as a Source of Data in Research With Older Adults. **Qualitative Health Research**, v. 15, p. 991-997, 2005.

JOHNSON, L. F. . L. A. . S. R. S. . &. H. K. Key emerging technologies for elementary and secondary education. **Education Digest**, v. 76, n. 1, p. 36-40, 2010.

JOHNSON, L. F. et al. Key emerging technologies for postsecondary education. **Education Digest**, 76, n. 2, 2010. 34–38.

KALKOFEN, D. et al. Visualization Techniques for Augmented Reality. In: FURHT, B. **Handbook of Augmented Reality**. [S.l.]: Springer Science+Business Media, 2011. p. 65-98.

KAMARAINEN, A. M. et al. EcoMOBILE: Integrating augmented reality and probeware with environmental education field trips. **Computers & Education**, v. 68, p. 545-556, 2013.

KATO, H.; BILLINGHURST, M. **"Marker Tracking and HMD Calibration for a Video-Based Augmented Reality Conferencing System**. 2nd IEEE and ACM International Workshop on Augmented Reality. San Francisco: IEEE Xplore Digital Library. 1999. p. 85-94.

KAUFMANN, H.; SCHMALSTIEG, D.; WAGNER, M. Construct3D: A Virtual Reality Application for Mathematics and Geometry Education. **Education and Information Technologies**, 5, n. 4, December 2000. 263-276.

KENSKI, V. M. **Educação e Tecnologias:** O Novo Ritmo da Informação. Campinas: Papirus, 2007.

KERAWALLA, L. et al. Making it real": exploring the potential of augmented reality for teaching primary school science. **Virtual Reality**, v. 10, n. 3, p. 163–174, 2006.

KLEWENO, C. P. et al. The virtual retinal display as a low-vision computer interface: A pilot study. **Journal of Rehabilitation Research and Development**, v. 38, n. 4, p. 431-442, July/August 2001.

KNIVETON, J.; LLANAS, A. **Bounce Now**. Oxford: Macmillan, 2011.

KOSTARAS, N. N.; XENOS, M. N. **Assessing the Usability of Augmented Reality Systems**. 13th Panhellenic Conference on Informatics. Corfu: [s.n.]. 2009. p. 197-201.

KURZ, D.; BENHIMANE, S. **Gravity-Aware Handheld Augmented Reality**. IEEE International Symposium on Mixed and Augmented Reality. Basel: IEEE. 2011. p. 111-120.

L. CHOICE. Logical choice letters alive. Disponivel em: <http://bit. ly/MdQf8i>. Acesso em: 19 December 2014.

L. M. KMET, E. A. **Standard Quality Assessment Criteria for Evaluating Primary Research Papers from a Variety of Fields**. HTA Initiative. [S.l.]. 2004.

LARSEN-FREEMAN, D. Total Physical Response. In: LARSEN-FREEMAN, D. **Techniques and Principles in Language Teaching**. [S.l.]: Oxford University Press, 2000.

LEE, K. Augmented Reality in Education and Training. **TechTrends**, v. 56, n. 2, p. 13-21, March/April 2012.

LEWIS, G. **The Internet and Young Learners (Resource Book for Teachers)**. [S.l.]: Oxford University Press, 2004.

LIU, T.-Y. A context-aware ubiquitous learning environment for language listening and speaking. **Journal of Computer Assisted Learning**, v. 25, n. 6, p. 515–527, 2009.

MERCURY INPRESS. Incredebooks. **Incredebooks**, 2015. Disponivel em: <http://www.incredebooks.com>. Acesso em: 3 June 2015.

MERRIAM, S. B. **Qualitative Research:** A guide to Design and Implementation. San Francisco: Jossey-Bass, 2009.

MILGRAM, P. et al. Augmented Reality: A class of displays on the reality-virtuality continuum. **Telemanipulator and Telepresence Technologies**, v. 2351, p. 282-292, 1994.

MILGRAM, P.; KISHINO, A. F. A taxonomy of mixed reality visual displays. **IEICE Transactions on Information and Systems**, v. 12, p. 1321-1329, 1994.

MINISTÉRIO DA EDUCAÇÃO. **Programa Mais Educação**, 2013. Disponivel em: <http://portal.mec.gov.br/index.php?option=com_content&view=article&id=16690&Itemid=1113>. Acesso em: 18 Janeiro 2015.

MISTRY P, M. P. C. L. **WUW – Wear Ur World – AWearable Gestural Interface**. CHI. Boston: ACM. 2009. p. 4-9.

MOHER, D. et al. Research Methods and Reporting: Preferred Reporting Items for Systematic Reviews and Meta-analyses: The PRISMA Statement. **BMJ**, 339, 8 August 2009. 332-336.

MOONEY, C. G. **Theories of Childhood**. St Paul: Redleaf Press, 2013.

OCULUS VR. Oculus Rift - Virtual Reality Headset for Immersive 3D Gaming | Oculus VR | Oculus Rift - Virtual Reality Headset for 3D Gaming. Disponivel em: <https://www.oculus.com/rift/>. Acesso em: 18 December 2014.

ORTLIPP, M. Keeping and Using Reflective Journals in the Qualitative Research Process. **The Qualitative Report**, v. 13, n. 4, p. 695-705, 2008.

PBS KIDS. Cyberchase Shape Quest Mobile Downloads | PBS KIDS. Disponivel em: <http://pbskids.org/apps/cyberchase-shape-quest.html#>. Acesso em: 18 December 2014.

PECK, S. **Child-child Discourse in Second-language Acquisition**. Rowley: Newbury House, 1978.

PECK, S. Developing Chidren's Listening and Speaking in ESL. In: CELCE-MURCIA, M. **Teaching English as a Second or Foreign Language**. Boston: Heinle Cengage Learning, 2001. p. 139-149.

PECK, S. Language Play in Child Second Language Acquisition. In: LARSEN-FREEMAN, D. **Discourse Analysis in Second Language Research**. Rowley: Newbury House, 1980.

PIAGET, J. **The Child and Reality:** Problems of Genetic Psychology. Tradução de Arnold Rosin. New York: Grossman Publishers, 1973.

PIAGET, J. **The Principles of Genetic Epistemology**. New York: Basic Books, 1972.

PINKLEY, D. **Our World**. Boston: National Geographic Learning, 2014.

QUEENSBERRY, W. Using the 5Es to Understand Users - Whitney Interactive Design. **Using the 5es to Understand Users**, 2011. Disponivel em: <http://wqusability.com/articles/getting-started.html>. Acesso em: 7 July 2015.

RABBI, I.; ULLAH, S. A Survey on Augmented Reality Challenges and Tracking. **Acta graphica**, v. 24, p. 29–46, 2013.

RADU, I. Augmented Reality in Education: a Meta- Review and Cross-Media Analysis. **Ubiquit Comput**, 14 January 2014.

RADU, I. **Why should my students use AR? a comparative review of the educational impacts of augmented-reality**. International Symposium on Mixed and Augmented Reality ISMAR. Atlanta: IEEE. 2012. p. 313–314.

REDONDO, E. et al. **Educational Qualitative Assessment of Augmented Reality Models and Digital Sketching applied to Urban Planning**. Proceeding TEEM '14 Proceedings of the Second International Conference on Technological Ecosystems for Enhancing Multiculturality. Salamanca: ACM. 2014. p. 447-454.

REDONDO, E. et al. **New Strategies Using Handheld Augmented Reality and Mobile Learning-teaching Methodologies, in Architecture and Building Engineering degrees**. 2013 International Conference on Virtual and Augmented Reality in Education. [S.l.]: Elsevier. 2013. p. 52-60.

REILY, V.; WARD, S. M. **Very Young Learners (Resource Book for Teachers)**. [S.l.]: Oxford University Press, 1997.

ROBERTO, R. A. **Desenvolvimento de Sistema de Realidade Aumentada Projetiva com Aplicação em Educação**. Univerdade Federal de Pernambuco. Recife. 2012.

ROBERTO, R. A. et al. A Dynamic Blocks Platform Based on Projective Augmented Reality and Tangible Interfaces for Educational Activities. **SBC Journal on 3D Interactive Systems**, 4, n. 2, 2013.

ROLLAND, J. P.; DAVIS, L.; BAILLOT, Y. A survey of tracking technology for virtual environments. In: CAUDELL, W. B. A. T. **Fundamentals of Wearable Computers and Augmented Reality**. Mahwah: CRC, 2001. p. 67-112.

SALMI, H.; KAASINEN, A.; KALLUNKI, V. Towards an Open Learning Environment via Augmented Reality (AR): visualising the invisible in science centres and school for teacher education. **Procedia - social and Behavioral Sciences**, v. 45, p. 284-295, 2012.

SALVADOR-HERRANZ, G. et al. **Manipulating Virtual Objects With Your Hands:** A Case Study on Applying Desktop Augmented Reality at the Primary School. 46th Hawaii International Conference on System Sciences. Grand Wailea: IEEE Computer Society. 2013. p. 31-38.

SANDOR, C. . O. A. . B. B. . A. F. S. **Immersive Mixed-Reality Configuration of Hybrid User Interfaces**. IEEE and ACM International Symposium on Mixed and Augmented Reality. Vienna: [s.n.]. 2005. p. 110-113.

SANTOS, M. E. C. et al. Augmented Reality Learning Experiences: Survey of Prototype Design and Evaluation. **IEEE Transactions on Learning Technologies**, 7, n. 1, January-March 2014. 38-56.

SCOTT, W. A.; YTREBERG, L. H. **Teaching English to Children**. London: Longman, 1990.

SHELTON, B.; HEDLEY, N. **Using augmented reality for teaching earth-sun relationships to undergraduate geography students**. The First IEEE International Augmented Reality Toolkit Workshop. Darmstadt: IEEE. 2002. p. 1-8.

SILVA, M. et al. **AR Jigsaw Puzzle:** Potencialidades de Uso da Realidade Aumentada no Ensino de Geografia. XXV Simpósio Brasileiro de Informática na Educação. Dourados: Sociedade Brasileira de Computação SBC. 2014. p. 194-203.

SILVA, M. et al. **Glassist:** Using Augmented Reality on Google Glass as an Aid to Classroom Management. SVR 2014. Salvador: SBC. 2014. p. 37-44.

SILVA, M. M. . R. R. A. . A. T. V. **Um estudo de aplicações de realidade aumentada para educação**. Workshop de Realidade Virtual e Aumentada – WRVA. Paranavaí: Sociedade Brasileira de Computação. 2012.

SILVA, M.; ROBERTO, R.; TEICHRIEB, V. **Evaluating an Educational System Based on Projective Augmented Reality**. XXIV Simpósio Brasileiro de Informática na Educação. Campinas: SBC. 2013. p. 214-223.

SILVA, M.; ROBERTO, R.; TEICHRIEB, V. **Evaluation of Augmented Reality Technology in the English Language Field**. Anais do XXVI Simpósio Brasileiro de Informática na Educação. Maceió: Sociedade Brasileira de Computação. 2015. Paper accepted for publication.

SOARES, M. **Letramento:** um tema em três gêneros. Belo Horizonte: Autêntica, 1998.

SQUIRE, K.; KLOPFER, E. Augmented Reality Simulations on Handheld Computers. **The journal of the learning sciences**, v. 16, n. 3, p. 371-413, 2007.

SWAN, J. E.; GABBARD, J. L. **Survey of User-Based Experimentation in Augmented Reality**. 1st International Conference on Virtual Reality. Las Vegas: [s.n.]. 2005. p. 22-27.

TORI, R. **Educação sem distância:** as tecnologias interativas na redução de distâncias em ensino e aprendizagem. São Paulo: Editora Senac São Paulo, 2010.

TORI, R.; KIRNER, C.; SISCOUTO, R. **Fundamentos e Tecnologia de Realidade Virtual e Aumentada**. Porto Alegre: Editora SBC Sociedade Brasileira de Computação, 2006.

TUM. mirracle, 2012. Disponivel em: <http://campar.cs.tum.edu/files/mirracle/>. Acesso em: 17 December 2014.

VALE, D.; FEUNTEUN, A. **Teaching Children English:** A Training Course for Teachers of English to Children. Cambridge: Cambridge University Press, 1995.

WAGNER, T. **The global achievement gap:** why even our best schools don't teach the new survival skills our children need - and what we can do about it. New York: Basic Books, 2008.

WEST WINDSOR PLAINSBORO REGIONAL SCHOOL DISTRICT. **A Vision for WW-P: Competencies for 21st Century Learners**. West Windsor Plainsboro Regional School District. West Windsor, p. 4. 2009. (11556646).

WU, H.-K. et al. Current Status, opportunities and challenges of augmented reality in education. **Computers & Education**, v. 62, p. 41-49, 1 March 2013.

YIN, R. K. **Case study research:** Design and methods. 4th. ed. Thousand Oaks: Sage, 2008.

ZABALA, A.; DA ROSA, E. F. **A prática educativa:** como ensinar. Porto Alegre: Artmed, 2007.

ZHOU, F.; DUH, H. B.-L.; BILLINGHURST, M. **Trends in Augmented Reality Tracking, Interaction and Display:** A Review of Ten Years of ISMAR. 7th IEEE/ACM International Symposium on Mixed and Augmented Reality. Cambridge: IEEE. 2008. p. 193 - 202.

APPENDIX A: LIST OF SELECTED STUDIES

ID	Author and Year	Title
S01	Salvador-Herranz, G. et al. (2013)	Manipulating Virtual Objects with Your Hands: A Case Study on Applying Desktop Augmented Reality at the Primary School
S02	Ahn, J.G et al. (2013)	Supporting augmented reality based children's play with pro-cam robot: three user perspectives
S03	Riera, A. S. et al (2012)	Developing an Augmented Reality Application in the Framework of Architecture Degree
S04	Arvanitis, T. N. et al. (2009)	Human factors and qualitative pedagogical evaluation of a mobile augmented reality system for science education used by learners with physical disabilities
S05	Redondo, E. et al. (2013)	New Strategies Using Handheld Augmented Reality and Mobile Learning-teaching Methodologies, in Architecture and Building Engineering degrees
S06	Chen, C. M; Tsai, Y.N. (2012)	Interactive augmented reality system for enhancing library instruction in elementary schools
S07	Furió, D. et al. (2013)	Evaluation of learning outcomes using an educational iPhone game vs. traditional game
S08	Kamarainen, A. M. et al. (2013)	EcoMOBILE: Integrating augmented reality and probeware with environmental education field trips
S09	Lin, T.J. et al. (2013)	An investigation of learners' collaborative knowledge construction performances and behavior patterns in an augmented reality simulation system
S10	Chang, K. E. (2014)	Development and behavioral pattern analysis of a mobile guide system with augmented reality for painting appreciation instruction in an art museum
S11	Echeverría, A. et al. (2012)	Exploring different technological platforms for supporting co-located collaborative games in the classroom
S12	Fonseca, D. et al. (2014)	Relationship between student profile, tool use, participation, and academic performance with the use of Augmented Reality technology for visualized architecture models
S13	Cai, S. et al.	A case study of Augmented Reality simulation system application in a

	(2014)	chemistry course
S14	Salmi, H. et al. (2012)	Towards an Open Learning Environment via Augmented Reality (AR): visualising the invisible in science centres and schools for teacher education

APPENDIX B: QUANTITATIVE CRITERIA

ID	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	Total
S01	2	2	2	2	n/a	n/a	n/a	2	1	2	2	1	2	2	0.909
S02	2	2	2	2	1	n/a	n/a	2	2	2	2	2	2	2	0.958
S03	2	2	2	2	1	n/a	n/a	2	2	2	2	2	2	2	0.958
S04	2	2	2	2	n/a	n/a	n/a	2	1	2	1	1	2	2	0.864
S05	2	2	2	2	2	n/a	n/a	2	2	2	2	2	2	2	1.000
S06	2	2	2	2	2	n/a	n/a	2	2	2	2	2	2	2	1.000
S07	2	2	2	2	2	n/a	n/a	2	2	2	2	2	2	2	1.000
S08	2	2	2	2	n/a	n/a	n/a	2	2	2	2	1	2	2	0.954
S09	2	2	2	2	1	n/a	n/a	2	2	2	2	2	2	2	0.958
S10	2	2	2	2	n/a	n/a	n/a	2	2	2	2	2	2	2	1.000
S11	2	2	2	2	2	n/a	n/a	2	2	2	2	2	2	2	1.000
S12	2	2	2	2	1	n/a	n/a	2	2	2	2	2	2	2	0.958
S13	2	2	2	2	2	n/a	n/a	2	2	2	2	1	2	2	0.958
S14	2	1	1	2	n/a	n/a	n/a	2	2	2	2	1	2	2	0.864

C1: Question/objective sufficiently described?
C2: Study design evident and appropriate?
C3: Method of subject/comparison group selection or source of information/input variables described and appropriate?
C4: Subject (and comparison group, if applicable) characteristics sufficiently described?
C5: If interventional and random allocation was possible, was it described?
C6: If interventional and blinding of investigators was possible, was it reported?
C7: If interventional and blinding of subjects was possible, was it reported?
C8: Outcome and (if applicable) exposure measure(s) well defined and robust to measurement/misclassification bias means of assessment reported?
C9: Sample size appropriate?
C10: Analytic methods described/justified and appropriate?
C11: Some estimate of variance is reported for the main results?
C12: Controlled for confounding?
C13: Results reported in sufficient detail?

C14: Conclusions supported by the results?

APPENDIX C: QUALITATIVE CRITERIA

ID	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	Total
S01	2	1	2	2	2	1	2	2	2	2	0.9
S02	2	2	2	2	2	2	2	2	2	2	1
S03	-	-	-	-	-	-	-	-	-	-	-
S04	2	2	2	2	2	2	2	2	2	2	1
S05	-	-	-	-	-	-	-	-	-	-	-
S06	-	-	-	-	-	-	-	-	-	-	-
S07	2	2	2	2	2	2	2	2	2	2	1
S08	2	2	2	2	2	2	2	2	2	2	1
S09	2	2	2	2	2	2	2	2	2	2	1
S10	2	2	2	2	2	2	2	2	2	2	1
S11	-	-	-	-	-	-	-	-	-	-	-
S12	-	-	-	-	-	-	-	-	-	-	-
S13	2	2	2	2	2	2	2	2	2	2	1
S14	-	-	-	-	-	-	-	-	-	-	-

C1: Question/objective sufficiently described?
C2: Study design evident and appropriate?
C3: Context for the study clear?
C4: Connection to a theoretical framework/wider body of knowledge?
C5: Sampling strategy described, relevant and justified?
C6: Data collection methods clearly described and systematic?
C7: Data analysis clearly described and systematic?
C8: Use of verification procedure(s) to establish credibility?
C9: Conclusions supported by the results?
C10: Reflexivity of the account?

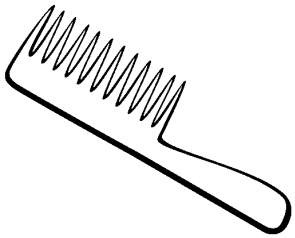
APPENDIX D: PILOT STUDY FINAL EVALUATION

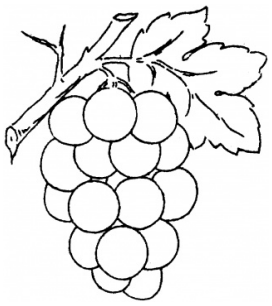
NOME: _____

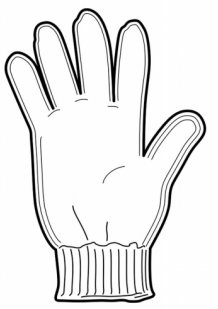
DATA: _____/_____/_____

TRABALHANDO COM RIMAS

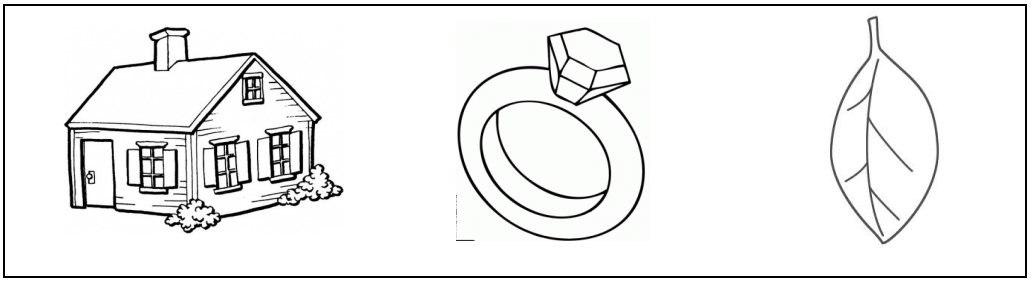
1) LIGUE AS FIGURAS QUE RIMAM:

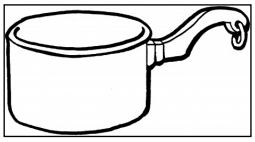


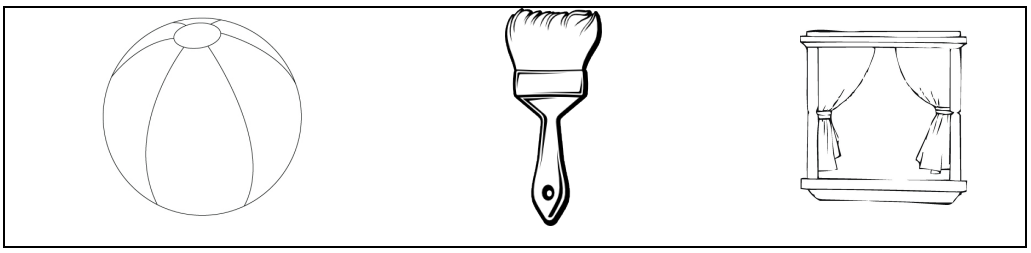


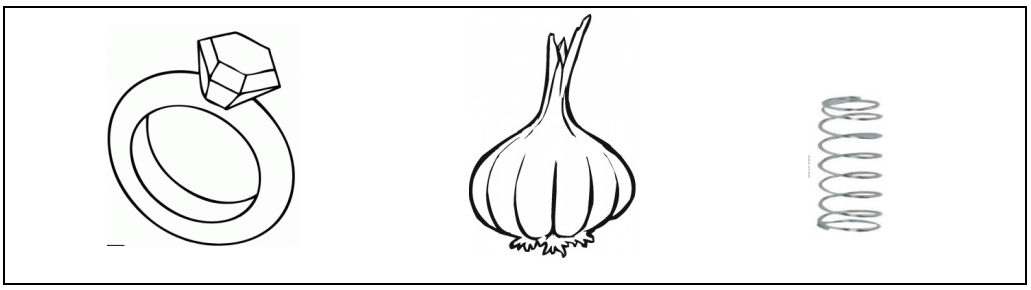


2) ENCONTRE E CIRCULE A FIGURA QUE RIMA COM AS FIGURINHAS EM DESTAQUE:









3) LEIA A PARLENDA E COMPLETE-A COM AS PALAVRINHAS ABAIXO:

MEIO ☐

MACACA ☐

☐ NO FOGO

BARRIGA ☐

☐ DIA

☐ ASSOBIA

FAZENDO ☐

PRA DONA ☐

DIA VAZIA

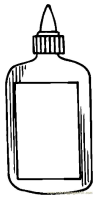
MACACA MEIO

CARETA VAZIA

PANELA	SOFIA

4) COMPLETE AS PALAVRAS COM AS VOGAIS QUE FALTAM:

B		L	



C		L	

APPENDIX E: RESEARCH DIARY PAGE

Classroom information

Track: _____

Time: _____

Date: ____/____ /_____

What kind of activities did you work today using the ARBlocks?

_____
_____
_____

What were the skills practiced using the application:

_____
_____
_____

How did the students interact with the ARBlocks?

( ) small groups ( ) the whole group together ( ) other

_____
_____

Did the teacher interact with the group during the activity?

( ) yes ( ) no

How did the students react to the activity?

_____
_____
_____

Did you observe students learning something using the application? If so, can you mention what was that?

_____
_____
_____

Did the students have any difficulties using the application? If so, can you mention them?

_____
_____
_____

Please, feel free to write any comment or situation you may find important today

_____
_____
_____
_____
_____
_____

APPENDIX F: PRE-KINDER 2 FINAL EVALUATION

Name:_____ Date: _____

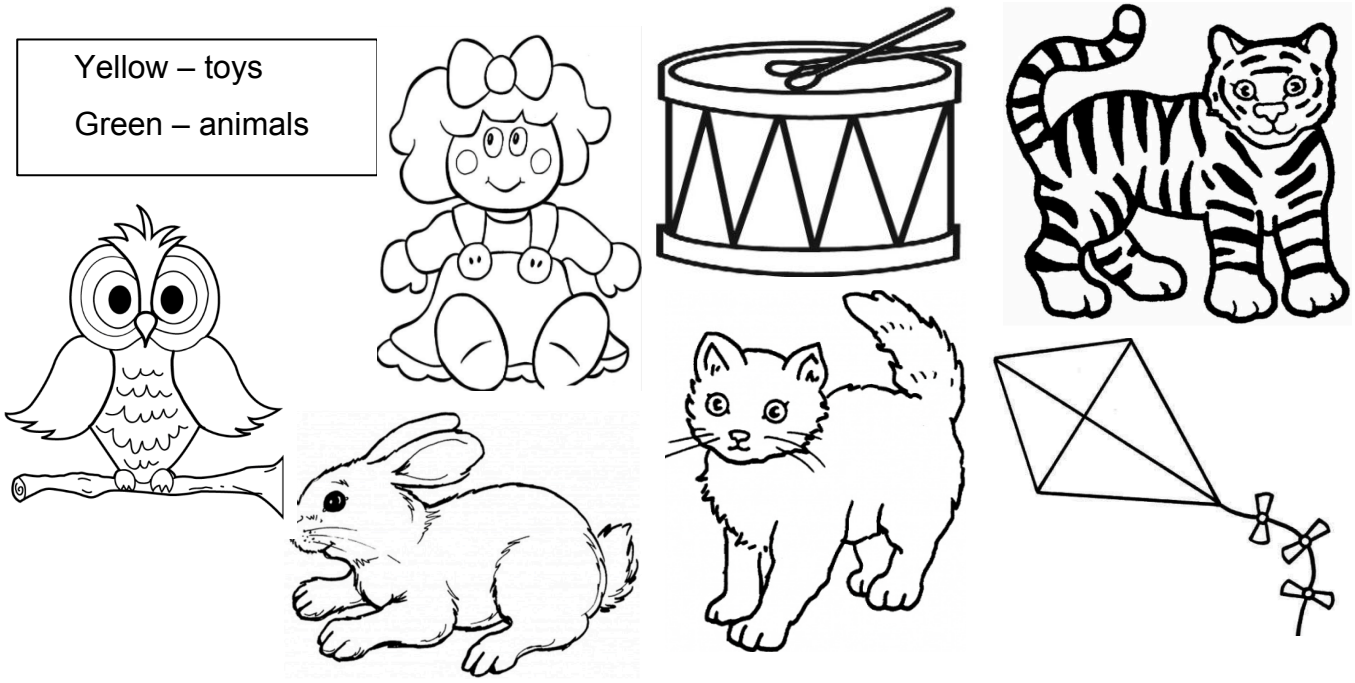
Activity Pre-kinder 2

1) Match the sentences

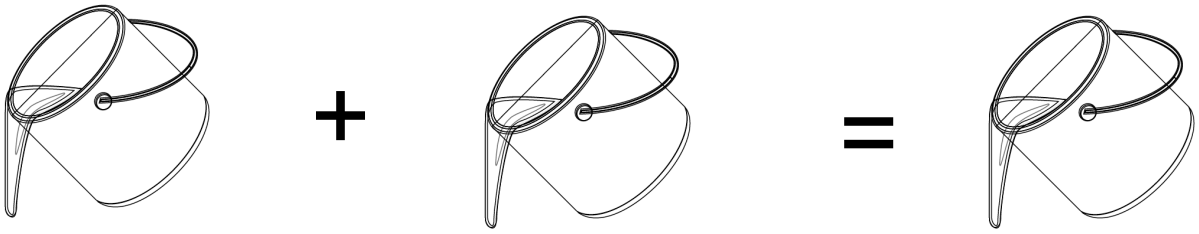
1. Hello.	Fine, thank you.
2. What's you name?	Hello.
3. How are you?	My name is _____.

2) Color the pictures according to the code in the box

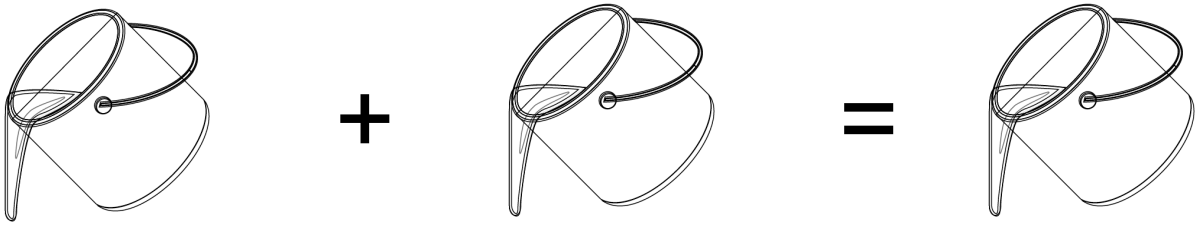
Yellow – toys
Green – animals



3) Color the drops and discover the colors we can do by mixing:

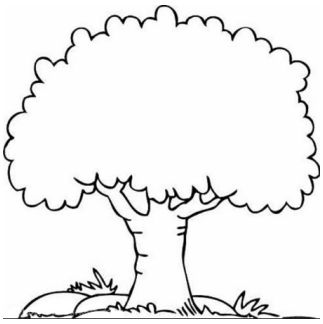


YELLOW + RED =



BLACK + WHITE =

4) Unscramble the letters:



E R

E T

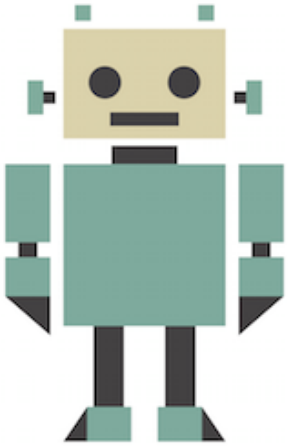
_____

R L O

F E W

_____

5) Count how many shapes are in the picture and answer the questions:

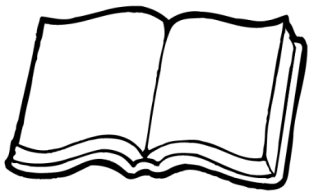


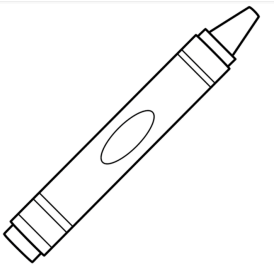
How many **triangle**s are in this picture?

How many **circles** are in this picture?

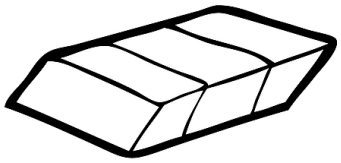
6) Match the pictures in the correct places:

a





a

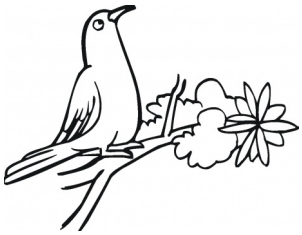


APPENDIX G: KIDS 1 FINAL EVALUATION

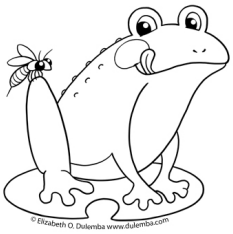
Name:_____Date: _____

1) Look at the pictures and choose the preposition that best describes it:

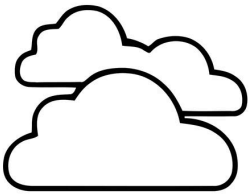
The bird is _____ the tree. (in/on)



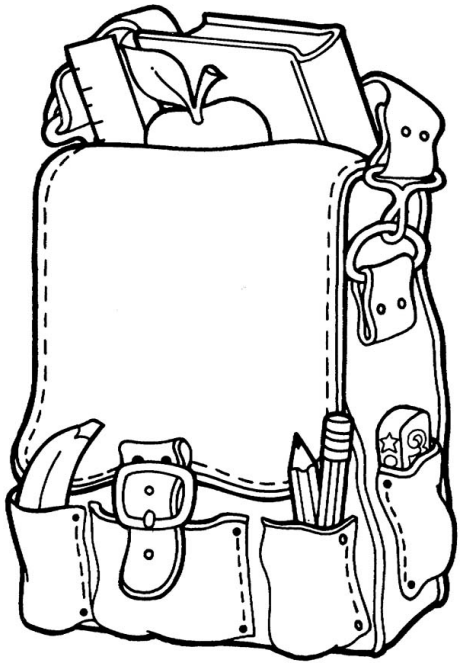
The frog is _____ the lily pad. (in/on)



The clouds are _____ the sky. (in/on)



2) Look at the picture and answer the questions.



a) how many pencils?
_____

b) how many books?
_____

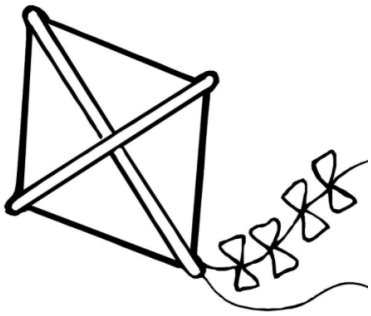
c) how many fruits?

3) Read the descriptions and draw.

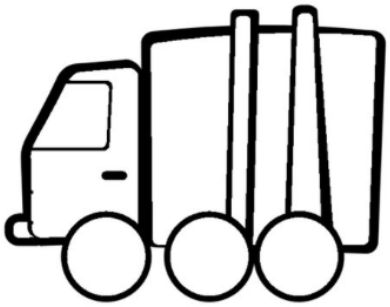
I am Lisa. I am wearing a pink dress and black sandals.

I am Bob. I am wearing a gray t-shirt, blue shorts and brown tennis shoes.

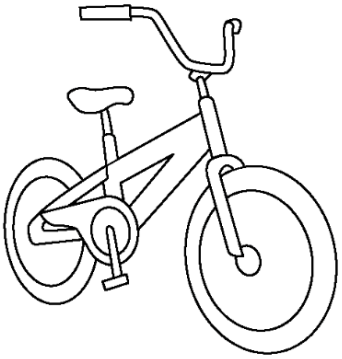
4) Match the pictures to its descriptions



kite



bike



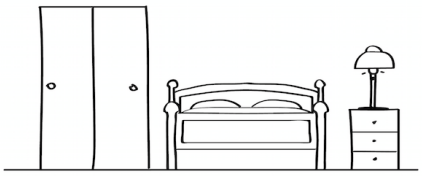
truck

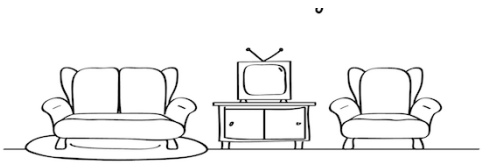
5) Match the pictures and the sentences.

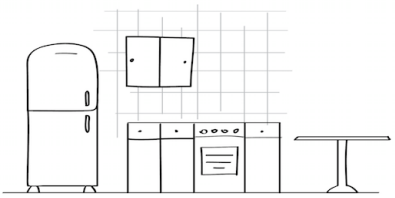
The armchair is in the living room.

The fridge is in the kitchen.

The bed is in the bedroom







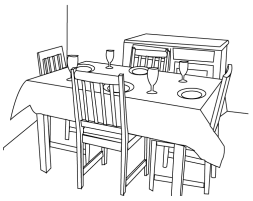
6) Unscramble the words to find the names

laasd

_____

kenci

_____



nigndi

_____

orombt

_____

APPENDIX H: QUESTIONNAIRE FOR THE PRE-KINDER 2 AND KIDS 1 STUDENTS

Name:_____ Date: _____

Reflection about the ABlocks

Pinte os quadrinhos de acordo com as carinhas

☺ verde

😐 amarelo

☹ vermelho

	☺	😐	☹
Eu gostei de usar os ARBlocks			
Eu gostei das atividades com os ARBlocks			
Eu aprendi inglês com os bloquinhos			
Foi fácil usar os bloquinhos			
Eu gostaria de brincar novamente com os bloquinhos			
Eu recomendaria essa brincadeira para um amigo			