Working Paper Series
ISSN 1177-777X

Using Augmented Reality to Enhance Children's Books

Nicholas Vanderschantz, Annika Hinze, and Aysha AL-Hashami

Working Paper: 02/2018
September 2018

Department of Computer Science
The University of Waikato
Private Bag 3105
Hamilton, New Zealand

Using Augmented Reality to Enhance Children's Books

Nicholas Vanderschantz[1], Annika Hinze[1], Aysha AL-Hashami[2]

[1,2] University of Waikato, Hamilton, New Zealand
[1]{vtwoz,hinze}@waikato.ac.nz [2]assha1@students.waikato.ac.nz

Abstract. This paper reports a case study on using Augmented Reality in children's books in which we explored the use of various types of interactions at different levels. The paper describes the design process and the insights gained through investigations into the requirements for enhancing children's books with Augmented Reality. Using evidence from the literature along with our own observations, an interaction design model and resulting interface design are developed to explore interactivity in printed books using Augmented Reality. We report on an expert walk-through and discuss possible improvements and implications from our case study.

Keywords: Augmented Reality, children's books, interaction model, reading.

1 Introduction

In this paper, we report on our case study into the opportunities of using Augmented Reality (AR) enhancement to printed children's picture books. The AR elements were developed by a design team following an established design process to identify the problem, hypothesize initial design solutions, iterate and develop these design solutions in preparation for testing. The findings presented in this paper will provide insight into the design and further testing of AR enhanced children's picture books.

The paper focusses on the design decisions and interactions and will not discuss technical details of the AR implementation process. We describe the design decisions, report on the proof-of-concept implementation, and discuss insights and implications. This paper extends the work reported in Vanderschantz et al. (2018) and provides a broader reporting of this investigation and the supporting expert walk through user study that space did not allow to be reported in that paper.

The paper is structured as follows: Section 2 discusses related work, while Section 3 introduces the material used for our case study. Section 4 introduces the three levels of interactivity created for the children's book and describes the prototype implementation. Section 5 summarizes the results of two expert evaluations and shows the final prototype. Section 6 discusses insights from the project and recommendations for future developments.

2 Related work

The potential of Augmented Reality (AR) integration on mobile devices has matured in recent years. AR has been explored for educational use (cf. Huang, Li and Fong, 2016) and AR books (cf. Cheng & Tsai, 2014) may assist children and parents during shared reading of children's books.

Interactive Books. It is well established that children may benefit from picture books through the consumption of both words and pictures (Bloom, 2002). The imagery incorporated in a children's book enables the comprehending of concrete ideas through the reinforcement provided by "nonverbal representations" (Piro, 2002). Pictures also provide an important support for children when the text is such that it is above the reading level of the child. Consequently, the appropriate design and composition of images on the page is essential to a successful, engaging children's picture book. Recommendations for how to use enhanced imagery and interaction in children's books is lacking from the literature. We previously studied interactive children's books (Timpany & Vanderschantz, 2012): physical interactions included pop-ups and opening additional pages to lifting flaps, tactile content or creating or re-arranging content, while intellectual interaction guided the reader in a non-linear course through content, requiring them to solve puzzles or make decisions to effect the outcome of the story. Most interactive books feature either physical interaction or intellectual interaction, and few books cater for both forms of interaction (Timpany & Vanderschantz, 2013). Presently eBooks provide video, audio and interactive elements and some interactive eBooks enhance engagement with the storyline through user in-book and out-of-book tasks. While digital books offer the opportunity for enhanced engagement with the storyline in ways that printed books cannot, however this interactive affordance is yet to be fully explored (Itzkovitch, 2012).

AR Books for learning. Cheng & Tsai (2014) studied the way children and parents read AR books and analyzed their behavior patterns and cognitive attainment. The study was based upon four categories of reading behaviors: *parent as dominator, child as dominator, communicative child-parent pair,* and *low communicative child–parent pair*. Cheng and Tsai found that the children in the *communicative child–parent pair* and *children as dominator* categories showed better learning outcomes, especially the children in the "communicative child–parent pair". Thus, children are actively involved in the process of AR books reading when they are with their parents, since they showed more cognitive attainment in terms of providing extensive and appearance descriptions. They suggest that AR books allowed the children to use their imagination.

Huang, Li and Fong (2016) observed a sample of 30 children (aged 4 to 5 years), including two teachers and two parents. The study was carried out by using a mobile application called *colAR Mix (*Puteko Limited, 2013). They found that using interactive tools such as AR technology has positive effects on learning, because it increases the level of creativity and imagination.

Yilmaz, Kucuk and Goktas (2017) studied the attitudes of preschool children toward AR picture books (ARPB) and examined the children's story comprehension performance (SCP), enjoyment/happiness during reading. They found a positive relationship

between ARPB-based learning and happiness: children who were happy with ARPB scored better marks on their story comprehension performance.

Flashcards are a popular tool for alphabet learning. An AR alphabet application was developed and studied by Rambli, Matcha and Sulaiman (2013). The application used a webcam to view the 3D objects and pattern markers and a computer. The researchers found that the children reacted positively toward the AR alphabet book. Most of the children requested to use AR technology in learning, and most of them were excited with the appearance of the 3D objects on the markers.

Educational magic toys. Besides books, toys are used as educational tools for early childhood education. AR-enhanced toys may create enjoyable learning opportunities for children or, conversely, distract the children (Yilmaz, 2016). The study used 'educational magic toys' (EMT) using Augmented Reality technology. Yilmaz found that children and teachers liked the EMT activities and they readily accepted it due to the 3D objects and animations linked with the toys. The teachers found the toys useful and easy to use. Although the children interacted and enjoyed the EMT, they reported a low level of cognitive attainment.

3 Study Material and Process

This section explores the materials that were used for the AR application and study.

We used the children's book *Hannah's Favorite Place* (Fiona Mason, 2015), a book written and illustrated specifically for the investigation of children's interactive picture books. The book has been developed as a traditional printed picture book as well as an adaptive printed children's book (Wright, 2015), see example pages in Figure 1.

Fig. 1. Illustrations of the tradition (L) and Physically Enhanced (R) book.

For this investigation, we selected a single spread (i.e., two pages) showing one of Hannah's favorite places: the zoo. The text on this spread reads: "Her auntie took Hannah on trips to the zoo; they'd look at the pandas, the chimps and birds too" with a supporting illustration, some of which is mentioned in the text (see Figure 2).

Fig. 2. Spread from the children's book *Hannah's Favorite Place* (Fiona Mason, 2015).

3.1 Study Approach: Design thinking

We summarize here our design considerations for the AR application:

Focus on objects mentioned in the story text. The objects to animate in the AR view were decided to be those that were mentioned or described in the text of the story. The decision for each object depended on whether it was a primary object on the story or not. For example, the text in our example page explicitly mentions the panda, so it should be animated and pop-up when scanning the page. On the other hand, the tree is not mentioned on the text, and therefore was not animated. The primary objects on the chosen page were: Auntie and Hannah, the pandas, chimps and birds (cf. Figure 2).

Determining objects to be visible or animated. While a number of in-animate objects can easily be shown on a page, fewer objects can be included in animations to avoid crowding on the phone screen. The animations may overlap and result in confusion. Furthermore, too many elements on a screen could reduce the ability of the camera to scan, recognize and display the correct animated object. Having fewer animated objects was also deemed positive to with regard to directing a child's attention and raise the level of concentration on a specific object.

Division of the target images. Each animation part must be unique and easy to recognize by the camera. It was decided not to encode any specific order for the animated objects to pop-up on the screen (but rather leave the selection to the reader). The division of the target image into parts was designed to make the virtual viewing experience easier to track. Having different parts also increases the ability image to be easily recognized by the camera, which then enables to animation.

Interaction types for different elements. It was decided to assign different types of interactions to the objects and to make these dependent on the element and the target reactions. That is, we selected objects to elicit different reactions, such as fun, easy learning, and learning. We used both visual animations (appearing objects, movement) and sound animations (animal sounds, spoken text). Simple appearance of objects was selected to increase the level of excitement for little children. It was designed to get them to pay more attention to the image, as they were waiting for exciting things to

happen. An example of this is to hide a bird from the constant image and display it on the phone screen when scanning a specific part of the image target.

Camera level and distance. While children may hold the camera closer to the page when scanning the image, parents may trial different distances to the page to explore various levels. We decided to therefore make the required level of the camera flexible in order to suit different users. Camera angle also need to be taken into account for correct identification of mages. In addition, a user could hold the phone horizontally, vertically or diagonally.

3.2 Study Design: Image Target

The image target that was used for the AR enhancement converted some of the page elements (see Section 3.1) into objects to be displayed in AR mode. Figure 3 shows the image target that was used in this project. This image target is similar as the spread that found in the original version of the *Hannah's favorite place* book, however some elements were initially hidden. Figure 2 shows the image as it first appears to a reader, while Figure 3 shows all animated objects. It can be seen that the following objects are initially hidden: the bird flying in front of the cat, the banana that the chimp was holding, as well as the drawing of the other chimp. These elements are discoverable through the use of AR interactions. To utilize and cover the whole image when scanning, it needs to be divided into levels, and the level of the camera when scanning the target image needs to be considered. Experiments were conducted to test all the possibilities of holding the camera by either the children or adults. We decided to divide the image into three levels as shown in Figure 3. This decision was reached through testing the angles and elevation levels of a phone's camera (with regards to the image) that had good focus, and also the goal that a camera could display at least one complete object.

Dividing the image target into three parts resulted in having nine image targets. Each level had different numbers of targets, and each target had different animation. The three levels were:

Level 1 (purple frame in Figure 3): The user holds the camera close to the image target so the camera scans one object at a time. (Thus, there are six image targets in this level).

Level 2 (red frame in Figure 3): The user holds the camera slightly further apart from the image target so the camera can scan about one-third of the image target, i.e. two objects at a time. At this level, there are two image targets, and different objects appear into each part.

Level 3 (yellow frame in Figure 3): The user holds the camera furthest away from the image target so the camera is able to scan the whole target at one time. There is only one image target to be discovered.

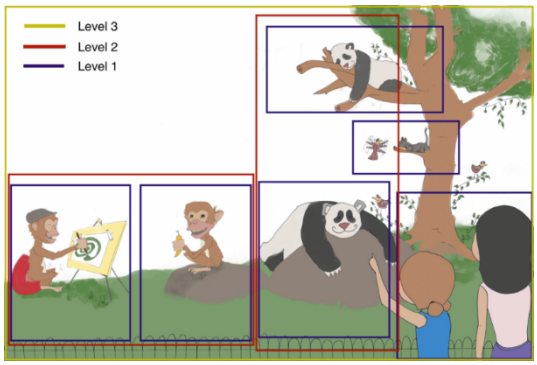

Fig. 3. The three levels of the image target.

3.3 Animation

The objects that will show up on the device screen will be objects of the characters that are in the image target at the selected level. Each object had a different animation, which depended on the object itself, and the way that the object was presented in the physical image and the story. Some objects were associated with sounds and simple movements, whereas some objects would be hidden from the physical/constant image and new objects would pop-up in their place. An example of an initially hidden object is the banana that one of the chimps is holding. When scanning the chimp, the banana will appear and the chimp will make arm movements as if eating the banana. When scanning the other chimp, the image the chimp is painting will appear as if drawn by the animal. Having different types of animations/movements and sounds were designed to help increase the level of excitement and allow the children to use their imagination.

4 The three levels of interactivity

The application has three levels of AR interactivity. These levels depend on how close the camera is when scanning the target. In each level, different animated objects combine both the enjoinment and learning.

Level one (simple fun interactions). In this level, the camera is closer to each object (assuming children will hold the phone). At this level, all objects were animated individually. A user can view and explore six primary objects (i.e. Hannah and her auntie, two pandas, two chimps, birds and a cat) of the target page, and each object will pop-up individually. Additionally, sounds will be played, such as screaming chimps when one of the chimps is being shown in the AR view.

This level is designed for fun and enjoying animations that appear in the level. The level of the animation in this stage is appropriate for children since it is simple and reflecting the current states of the objects. Figure 4 is a screenshot of the mobile app that shows an object from level one.

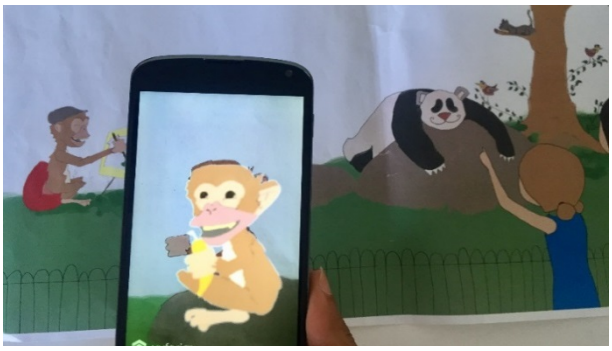

Fig. 4. An object that appears in level one.

Level two (simple educational interactions). The camera in this level is slightly away from the objects, scanning two objects at a time. This level is educational but in a fun way since children can learn how to spell words and recognize the related object of that name. As shown in the phone screen in Figure 5, this level introduces words and arrows that point to existing objects in the target page.

Users can explore two different animated texts and objects in this level, which are the two chimps and the two pandas. For example, the text that appears in this level is 'There are two chimps', and this sentence will show up word by word. Then a first arrow will point at the first chimp and a second arrow will point at the second chimp. After that, the word 'CHIMP' will appears letter by letter.

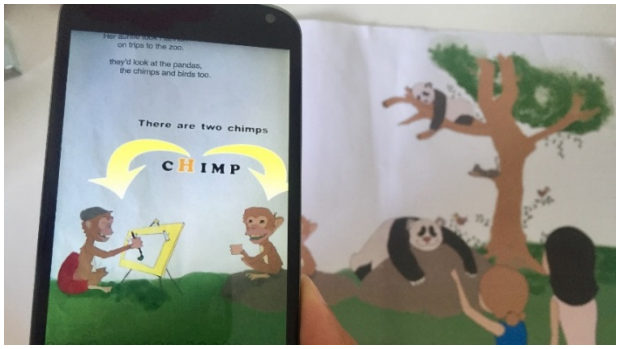

Fig. 5. Shows the word and arrows that appear in level two.

Level three (learning interactions). The last level allows the scan of the whole target page since the camera is away from the target page. This level was implemented to cover both the enjoyment and learning aspects. Children can explore new objects in this level, which are the bees that appear only at this level.

Moreover, in level three the story text was played in the background, i.e. the children can hear a person reading the text for them. Introducing new objects in this level could increase the excitement of the children and motivate them to hear the attached audio. Figure 6 is a screenshot that shows the bees on the third level.

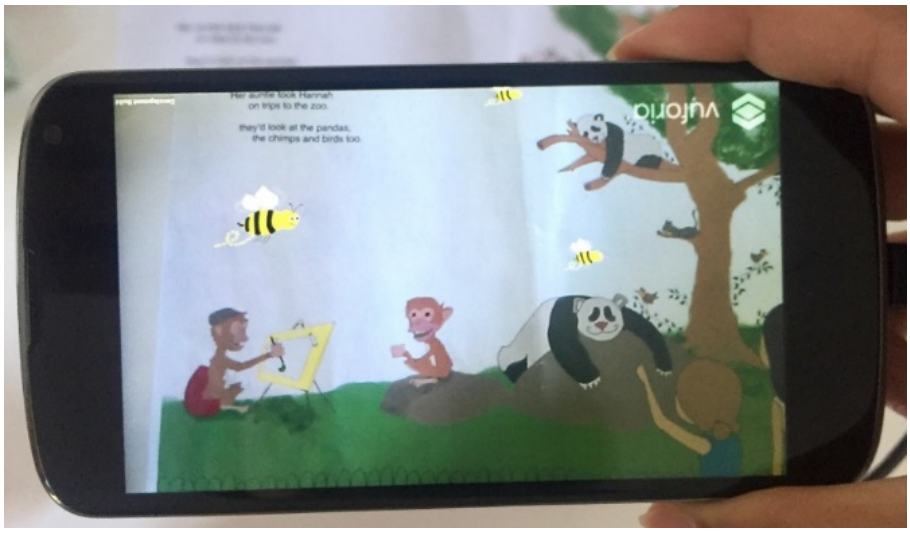

Fig. 6. Screenshot shows the bees in level three of the application.

5 Expert Walkthrough

In order to evaluate the software experience, an expert/cognitive walkthrough technique was used (Helander, Landauer & Prabhu, 1997). The expert walkthrough evaluation was performed twice, each with a different focus. The first evaluation was to examine the appropriateness of the initial software, determine possible problems, and provide suggestions to improve the software. The second expert walkthrough evaluation was done after the enhancement of the first software and took into account the issues raised in the first assessment.

5.1 Purpose & Procedure

Two evaluation sessions were held to assess the app at different implementation stages.

First evaluation session. The first expert evaluation session was conducted to obtain feedback about development of first draft of the application. The application in this testing stage was not fully implemented, and only its first level was ready to be tested. Another purpose of running this first session was to investigate and understand the users' behaviors and thinking when using the app. knowing the way that users will extract and deal with the application will help to improve the development of the app.

Second evaluation session. The second evaluation session was organized to evaluate the software application after enhancing the first draft according to the feedback of the first evaluation session. In addition, in this session the software application was fully developed, i.e. the three different levels were working as intended. Thus, the objective of this evaluation session was to get the final comments and feedback for the application from the experts - participants- who examined the app in the first session.

Additionally, this session gave the opportunity to conclude and state recommendations for any further future work as it was the last session that involved users of such apps.

5.2 Participants

For the first evaluation session, two participants, a male (EP1) and a female (EP2), conducted the walkthrough evaluation for the AR picture book reading app. For the second session, the same two initial experts took part again, and were joined by another female expert (EP3) joined them. All participants are experts in design, mobile technologies, and AR applications that are relevant to children.

When the walkthrough evaluation finished, each participant was interviewed to reflect their experience while using AR technology with a single page from a children's storybook. The primary objective was to investigate and discuss whether this application could be useful and educational for 4-6 years old's. The main difference between the two sessions was that the second session was an evaluation for the completed proof of concept application containing three levels of interactivity. Additionally, in this session, a guide sheet, intended as the first page of a book, was provided to the participants that explained the three levels of scanning, and the number of objects that can be viewed in each level.

5.3 Evaluation One: Results

While the two experts were exploring the application and the spread, their reactions ranged from excitement to disappointment. Both EP1 and EP2 were excited when they experienced the pop-up AR objects on the screen (e.g., the panda started to snore, and the chimp began to scream). Both exports reported frustration when trying to select objects at the different levels: it was unclear to them how many objects they were supposed to see/scan. EP2 stated that when scanning the bird, nothing happened although it was mentioned in the text, so a user will assume that he/she might find something when scanning the bird. In addition, when EP1 and EP2 noticed that the screen seemed too small to scan some objects. Both experts found it hard to determine how close the camera should be from the book, i.e. the height of the camera from target objects on the page. Furthermore, it was hard to share the screen between two people. Both EP1 and EP2 wondered if the experience would be different on a tablet.

From the first evaluation session, both EP1 and EP2 agreed that the enhancements would be fun and enjoyable for 4-6 years old children. EP1 stated that both illustration and animation level along with the sounds attached to the animation, was playful for that age group, especially that the animation was of their level and did not overtake whole story. EP2 said that there was a clear differentiation between reading material and illustration materials and the AR enhancements did not interfere with the ability to read since AR could separate the two materials from each other (reading materials and enhancements).

EP2 thought that the application could be more educational, suggesting that sounds, for example, could be appropriate for extending the educational affordances when

working AR for children's books. EP1 thought that the application did show some levels of education such as teaching children how to read and link the words. Both experts agreed that the movement and animation appropriately reflected the story. EP2 suggested coming up with a way that guides the parents to an appropriate sequence to follow to scan the spread. The participant thinks that helping the parents to explore the spread by providing them with some hints can increase the excitement of the children, since the self-exploration may end up not as exciting as if the parent follows a guided sequence.

5.4 Evaluation Two: Results

This evaluation session was with three experts. EP1 commented on the usefulness of the guide-sheet that was used to introduce the app during the evaluation session, and suggested adding the sheet in the front of the book. This version of the app included a number of additional interactive items and the experts enjoyed the range of interactions they discovered. It was discussed that this would ensure several revisits to the book which might offer multiple readings and multiple new discoveries for readers. Additionally this feature may reinforce the sharing between parents and their children, i.e. sharing their discoveries of new objects.

EP3 thought that the enhancements would be fun and enjoyable for 4-6 years old children, especially the newly included sounds which seemed to bring these enhancements to life. Additionally, EP3 suggested these kinds of enhancements might be suitable for even younger children (suggesting 2-3 years old's) because these were simple and attractive. The potential of expanding this project was discussed by EP3 who suggested that adding different features that serve different age groups would be valuable for such educational application. EP3 said that from a parent's point of view, children could use the application by themselves if they were assisted at the first time since children can pick things very quickly.

EP1 stated that level one of the application was more fun and enjoyable, whereas level two and three are more educational. EP1 also noted that highlighting the chimps, pandas and the bird by showing them on level one reinforced the story text, thus it could be said that this level is of an educational nature. In level two, EP1 said, the children could recognize the object when the arrows are pointing to them, and they learn how to spell the name of that object. For level three, EP1 suggested reinforcing the text when the audio played in that level to connect what was going on at this level.

5.5 Findings from both evaluations

It was recommended by the experts that a descriptive setup page be placed at the beginning of the AR book. Such a page might discuss ways of holding the mobile device, recommend device sizes if appropriate, and detail the use levels of interaction and how close the camera should be held.

In a spread that has multiple objects, indication points must be added to inform the user about the objects that can be scanned. Having these indications will reduce the

amount of time that the users may consume when searching for the possible places to scan to find any objects.

The size of the screen of the smart device should be appropriate for sharing. The size of the screen of the mobile device that was used in the evaluation session was not quite appropriate because it was very difficult for two users to view the display at the same time. Thus, having a device with a bigger screen size, i.e. tablets, could improve the experience of sharing the display at the same time.

Similar applications could work for different age group. This kind of educational application could work for different age of group such as children of 2-3 years old. The application could recognize the same target image but reacted slightly different in terms of the enhancements and sounds based on the level of the age group.

The integration of AR may assist with ensuring the usefulness and enjoyability of the book for a longer-term by adding different choices, options, and interfaces with a variety of educational and enjoyable levels that serve different age of groups. In line with this, perhaps such educational applications and indeed books should state their intended target group. Identifying the likely appropriate age of the children who could use the application along with their ability or developmental stage might assist parents and teachers with making decisions about new technologies and books such as this.

5.6 Final AR application & interactions with book

Figure 7 shows examples of level one interactions with the book; Figure 8 shows level two interactions. Note the different AR animations on the same objects depending on level. (panda snoring on level one, Figure 7 left, and panda text on level two, Figure 8, right) The level three interactions are shown in Figure 6.

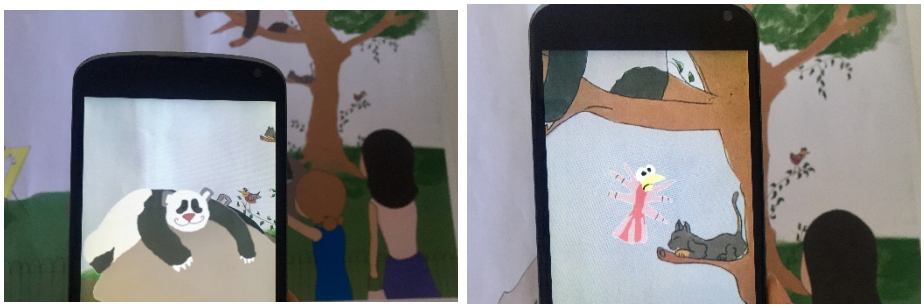

Fig 7: Level one (animated painting, appearing bird, snoring panda, bird sounds)

Fig 8: Level two (animated words exploring objects)

6 Conclusions & Recommendations

Augmented Reality (AR) is a technology that converts a real environment into a digital environment by displaying 3D or 2D models in the real environment. This technology could have a significant impact on future education. Cheng and Tsai (2014) found that children showed more cognitive attainment when they used an AR book together with their parents. AR technology has the potential improve children's learning through motivational mechanisms, intellectual interactions, and increased memorability of book content through the addition of unexpected animation and playfulness.

The aim of our project was to develop an AR application for a children's picture book that explored the potential to engage both physical and intellectual interactions. We used three levels of interactivity to provide a variety of interactions and engagement opportunities. Our application combines fun and learning aspects at different camera levels. We engage the user in a range of physical and intellectual interactions through different types of animation and user interaction with the page. We used sounds, animated objects, as well as appearing and disappearing objects which we believe will draw children's attention and delight. Revealing objects like the banana and the bird in the application aims to increase children's engagement. Adding new objects (such as bees and additional text) in the augmented animations attempts to encourage children to explore the target page further. Using a number of enhancements in the AR version of the book should appeal to children's sense of curiosity, encouraging children to compare the two versions of the book (physical book and AR application), and afford opportunities to explore the book in more detail and through multiple sittings.

6.1 Recommendations

To develop an educational AR application for a children's storybook, we recommend taking several considerations into account.

The application should **combine both playful interactions and learning focused interactions** to ensure attractiveness to a broad range of children at various reading ages and motivational levels. We showed that it is possible to replicate and extend the enhancements of a physical children's picture book explored by Wright (2015) using AR. Investigation of how to extend more of Wrights recommendations will benefit this investigation to discover the affordances of AR that are not present in printed three-dimensional picture books.

Using a **number of enhancements** in the AR version of the book is useful for attracting children's sense of curiosity: sound, movement, hiding object and adding new objects or games. This will encourage children to compare the two versions of the book (physical book and AR application), and thus attracts the children's focus. Adding completely new objects that appear only in the application were hypothesized to encourage children to explore more. It was noted that limits are present with how much is possible on the small screen compared to the flat larger page and therefore avoiding crowding is a consideration for the designer.

Using **multiple levels of camera targeting** allows for targeting several age groups in the same book and encourages exploration as well as co-reading/shared-reading with siblings and parents.

Further exploration of the possibilities of AR enhanced picture books at multiple interaction levels is warranted through user studies and development of further testing material to extend this course of investigation.

References

1. Bloom, P.: How Children Learn the Meanings of Words. MIT Press, Cambridge, MA, USA (2002).
2. Cheng, K.-H., Tsai, C.-C.: Children and parents reading of an augmented reality picture book: Analyses of behavioral patterns and cognitive attainment. Computers and Education, 72, 302-312 (2012).
3. Huang, Y., Li, H., & Fong, R.: Using augmented reality in early art education: A case study in Hong Kong kindergarten. Early Child Development and Care, 186(6), 879-894 (2016).
4. Itzkovitch, A.: Interactive eBook Apps: The Reinvention of Reading and Interactivity, http://uxmag.com/articles/interactive-ebook-apps-the-reinvention-of-reading-andinteractivity, (2012).
5. Piro, J.M.: The picture of reading: Deriving meaning in literacy through image. Read. Teach. 56, (2002).
6. Rambli, D., Matcha, W., & Sulaiman, S.: Fun learning with AR alphabet book for preschool children. Procedia Computer Science, 25, 211-219 (2013).
7. Timpany, C., Vanderschantz, N.: A Categorisation Structure for Interactive Children's Books. Int. J. Book. 9, 97–110 (2012).
8. Timpany, C., Vanderschantz, N.: Using a Categorisation Structure to Understand Interaction in Children's Books. Int. J. Book. 10, 29–44 (2013).
9. Vanderschantz, N., Hinze, A., & AL-Hashami, A.: Multiple Level Enhancement of Children's Picture Books with Augmented Reality. In International Conference on Asian Digital Libraries (in print). Springer, Cham. (2018).

10. Yilmaz, R. M.: Educational magic toys developed with augmented reality technology for early childhood education. Computers in Human Behavior, 54, 240-248 (2016).
11. Yilmaz, R. M., Kucuk, S., & Goktas, Y.: Are augmented reality picture books magic or real for preschool children aged five to six. British Journal of Educational Technology, 48(3), 824-841 (2017).
12. Wright, K.; A comparison of children's books: Picture books versus physically and intellectually adaptive children's books. Unpublished Master's Thesis, University of Waikato, New Zealand (2015).