Universiteit Utrecht

Master Thesis Game and Media Technology

INFOMGMT1

ICA-3591166

Augmented reality for cultural heritage - Comparing a cardboard device with a mobile handheld device in a museum setting

Submitted To:

Supervisor & first examiner:
dr. W.O. Hürst
Department of Information and Computing Sciences
Division Interaction Technology

Second examiner:
prof. dr. R.C. Veltkamp
Department of Information and Computing Sciences
Division Interaction Technology

Submitted By :
Remco de Koeijer (3591166)

Abstract

There are different options to implement augmented reality (AR), which range from high-end expensive solutions, to cheap, less sophisticated, alternatives. Some of those cheap alternatives are a mobile phone or a cardboard-based solution. A mobile phone creates so-called handheld AR, where users point their device to a point of interest and the video on the screen is augmented with virtual 3D objects. A cardboard solution makes use of a mobile phone to project virtual images on a see-through mirror, showing the images in the real world. Both these options use a mobile phone, but they create different experiences. One is handheld and shows a video, whereas the other is head-worn and projects the images in the real world. However, there is no comparative evaluation between these two alternatives that shows the differences between them. The goal of this thesis is to identify and evaluate these differences in the context of cultural heritage and find the better suited device for that situation. Both AR solutions utilize a state-of-the-art mobile phone and are compared based on the visitor experience. A user study is conducted with the visitors of a museum in Utrecht: the Rietveld Schröder House, a cultural heritage site which is part of the UNESCO world heritage list. Two prototypes are implemented, one for each AR solution, and tested with 28 visitors of the Rietveld Schröder House during regular museum tours. Both implementations showed good results, but each has its own advantages. The results from a standardized questionnaire on visitor experience used after the experiment shows a preference for cardboard AR, whereas the observations during the test indicate that handheld AR might be easier and more intuitive to use.

Disclaimer

This research is conducted as part of an internship with Capgemini. Capgemini is an international IT Consulting company and has a division aimed at innovation and finding business cases for new and rising technologies, such as AR. One of the partners of Capgemini is the Centraal Museum in Utrecht and they asked for an implementation of AR in the Rietveld Schröder House (RSH), which is owned by them. Their aim is to see how the RSH can benefit from the addition of AR and if it can enhance the regular museum tour. Due to copyright reasons, there are no images or photographs from the RSH or related to the RSH in this thesis. The internship is done together with Joost Overbeek, an industrial design student. He contributes to parts of this thesis by providing, for example, the design side of the project. His concrete input is mentioned throughout the thesis. All other work is done by the author.

Contents

1 Introduction

Augmented reality (AR) is a technology that is becoming more popular over the past few years. Only a few consumers know what AR is in 2016, but now many of them want to use AR to customize and personalize the world around them. For example, a recent study showed that over 50% of the consumers would like to have AR glasses that light up dark places and almost 40% would like to pick up digital items related to games or edit out shop windows and billboards that they are not interested in [1].

Looking at current AR usage, one sector that AR is used and experimented with is cultural heritage. It is shown that the addition of AR to cultural heritage allows, for example, visitors of a museum or cultural heritage site to have a better experience by learning more and having a more interactive visit. Two examples of such successful AR applications for cultural heritage are the ArtLens 2.0 discussed by Ding [2] and the explorative system for outdoor points of interest (POI) discussed by Brancati et al.[3]. The ArtLens 2.0 application gives visitors extra information about the art that they are viewing and helps to guide them to other pieces based on their personal interests. The explorative system helps to navigate users towards a POI of their selection and also gives extra information about that. Both applications have received positive feedback from the users, proving the statement that AR has a positive impact on visitor experience in this domain.

AR concepts

In general AR works by detecting and recognizing a 'marker' in the real world. This marker can be either a surface, a building, a location, an image, an object or something else. After the recognition, a virtual layer is added onto the real world, which can be a virtual heads-up display (HUD) or an object that is added to the real world on a 3D position. According to Azuma an AR system has the following three characteristics [14]:

- AR combines the real world and the virtual world.
- AR is interactive in real time.
- AR is registered in three dimensions.

Virtual objects placed in the real world are anchored to a position. With those virtual objects anchored, the user can walk around it and view it from up close. In some cases it is possible to interact with the virtual object. This is applied in the Skin and Bones app discussed in the Related Work section. See Figure 5. This follows the characteristics that are proposed by Azuma.

The HUD is a virtual overlay on the real world and provides the user with information on the object or location that is viewed. In some cases the user can interact with the HUD as if it is a user interface. The application for the Taipei Fine Arts museum discussed in the Related work section applies this. See Figure 6. The virtual overlay is 2D, not following the characteristic from Azuma. However, the HUD is only seen when a marker is detected which has a 3D position in the real world.

2 Research goal

This section looks at what AR offers to cultural heritage and what kind of systems are used to achieve this. Some practical considerations for implementing AR at a cultural heritage site are given and with those in mind, an alternative AR system is introduced. There is no comparative research between the given (alternative) AR systems and the research question on that is introduced.

2.1 Research context

Two major device categories for AR are mobile handhelds, such as smartphones or tablets, and head-mounted-displays (HMD's), such as a Hololens or a Google Glass [13].

A mobile handheld shows the real world seen through the camera of the device with the added augmentations on the screen of the device as a video. Some advantages of mobile handhelds are that they are intuitive to use, widely available, cheap and the augmentations are seen clearly on the screen of the device. Some of the disadvantages are that the user only sees a video on a screen, creating a 'flat' experience. The user also needs at least one hand to hold the device. An example for mobile handheld AR is the ArtLens 2.0 application [2] mentioned in the introduction.

An HMD allows the user to see the real world and the augmentations are projected on a see-through screen which is in front of the eyes of the user. An advantage of an HMD is that the objects actually seem to be in the real world, improving the immersion of the user. Another advantage is that the user has two free hands to do other things. With the see-through display, the augmentations are semi-transparent. If the user is in a bright environment, the augmentations are hard to see, a disadvantage of an HMD. Other disadvantages are that an HMD can be heavy and uncomfortable to wear and they are expensive. An example application from cultural heritage of an HMD is the explorative POI system [3].

For an AR implementation at a concrete cultural heritage site, several practical considerations have to be taken into account when deciding on what AR device to use. One consideration is that the device should be easy to use, as visitors should be able to use the application without the need for training or long instructions. Another consideration is that the device should not be cumbersome to use, as it should improve the visit of cultural heritage. Furthermore, the device should not be expensive, since many might be needed to supply all the visitors with a device.

2.2 Problem addressed in this thesis

Allowing many visitors the usage of an AR device at a cultural heritage site, the price of the device should be low. With a low price, more devices can be purchased. Considering the discussed AR devices, only the mobile handheld has a low price. An alternative to the expensive HMD's is cardboard AR. Companies like Aryzon [17], HoloKit [18] and Lenovo [19] are investigating this cheap AR HMD alternative.

Aryzon and HoloKit both produce similar AR headsets with a price that is around 30 euros. They produce an AR cardboard that uses a smartphone for all the technical aspects, such as detection and creating the images for projection. The created images are projected on a see-through mirror, through which the user can see the real world. This way of AR is close to what an HMD offers. The design of the cardboard of both companies is different and they use a different underlying AR software development kit. See Figure 1 and Figure 2 to see the Aryzon and Holokit cardboard respectively.

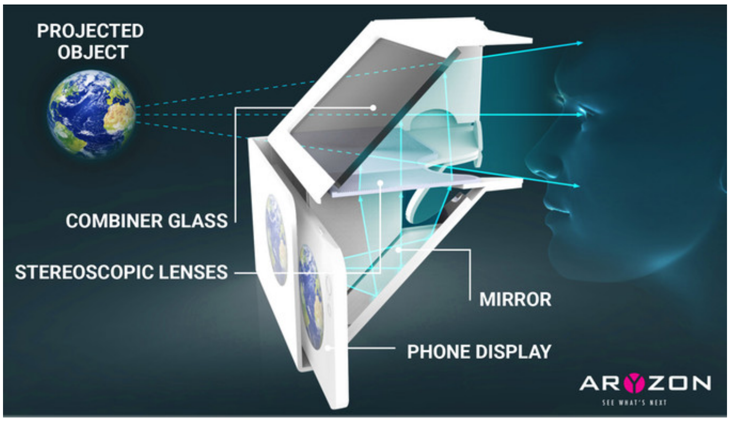


Figure 1: An overview of the Aryzon cardboard. Source: Aryzon Kickstarter page.

Figure 2: An image of the Holokit cardboard. Source: Holokit website.

The Lenovo Mirage is a bit more advanced. The headset is made of plastic and has built-in cameras for the detection. The headset can detect two detection beacons, one that is held by the user and one that is placed on the ground. A smartphone is placed inside the headset that creates the images for the projection. For pro-longed use, the phone can be attached to a charger cable that is inside the headset. Since the Lenovo Mirage is more advanced, the price is also higher compared to the Aryzon and HoloKit, at a price of 200 dollars in the United States or 250 euros in the Netherlands. See Figure 3 for an image of the Lenovo Mirage.

Figure 3: An image of the Lenovo Mirage. This includes a headset, a beacon as target and an object to be held in the hand that includes a tracker. Source: Lenovo website.

Cardboard AR, like the Aryzon and the Holokit, and the mobile handheld are both similar in price, but offer different ways of AR. Where cardboard AR projects the digital images on a see-through mirror, handheld AR shows a live video of the real world with the digital objects placed on it. Also, cardboard AR is worn on the head, which might be uncomfortable to the user. With mobile handheld AR, the user points the device to a target position, which might be uncomfortable in the arms. These differences might have an impact on the user experience and research is needed to identify that. This research can also determine the feasibility of using these AR devices in a real-world setting.

This research will look at which of the cheap AR solutions is the preferred choice in the context of cultural heritage. To narrow it down to a singular comparison, both categories are represented by one representative device. For mobile handheld AR the Samsung Galaxy S7 is chosen. This state-of-the-art device is capable of smoothly running an AR application. The cardboard AR category is represented by the Aryzon, as it was available earlier than the Holokit. The mobile phone for the Aryzon is the Samsung Galaxy S7.

2.3 Research question

This research looks at two cheap AR solutions, the Samsung Galaxy S7 and the Aryzon AR cardboard, to see which is preferred by visitors of the Rietveld Schröder House, a museum in Utrecht. To find the preferred device, the visitor experience is evaluated, which is split into three major factors: visitor enjoyment, visitor satisfaction of their visit and the knowledge gained by the visitor [5]. This results in the following research question:

What is the difference in experience (measured in enjoyment, satisfaction about the visit and knowledge gained) between an augmented reality cardboard headset (the Aryzon cardboard with standard smartphone) and an augmented reality mobile handheld solution (a standard smartphone) when visiting a cultural heritage building (the Rietveld Schröder House)?

2.4 Hypothesis

There are two categories in which both AR devices differ: visualization of AR and usability. The Aryzon projects the digital images on a see-through mirror, whereas the Galaxy S7 shows a live video. With the digital images seen in the real world with the Aryzon, the immersion should be greater compared to the Galaxy S7, leading to a better experience. However, realizing this technically is harder, making performance and implementation issues more likely. The expectation is that if there are no technical issues, the Aryzon creates the best visitor experience. Regarding the usability, this is split into ergonomic problems and the use of the AR device itself. Since the Aryzon is worn on the head, it might be cumbersome to use, whereas the Galaxy S7 has to be pointed at a target, which might lead to tiring or uncomfortable situations. If users are comfortable whilst wearing the Aryzon, the expectation is that it creates the better experience. Interacting with an application on the Aryzon is possible with, for example, hand gestures, making interaction harder. With the Galaxy S7 interaction is intuitive, making it easier to use. If there is no interaction possible with the application, the Aryzon has a better experience.

Overall, the expectation is that the Aryzon creates a better visitor experience, due to the immersion, if there are no big technical, ergonomic or usage issues.

3 Related Work

This literature review looks at some recent implementations of AR for cultural heritage. These papers are reviewed to get an insight on what kind of AR applications there are, how they are received by the visitors and what kind of research is conducted. This information is used as an inspiration for the development of the AR applications for this research. Papers from AR in other domains are reviewed as well, to find advantages and disadvantages that can be applied to AR in all domains and the AR applications for this research.

3.1 Mobile Handheld AR for cultural heritage

The AR device that is seen the most for cultural heritage is a mobile handheld, such as a smartphone or tablet. This is mainly due to the fact that mobile handhelds are more accessible and cheaper. Some recent mobile handheld AR applications for cultural heritage are reviewed by Ding [2] and discussed below.

ArtLens 2.0

The first application is ArtLens 2.0, an application for the Cleveland Museum of Art. This application for Android and iOS can be downloaded in less than a minute and enhances a visit with a personal tour and information about the artwork. It is capable of recognizing a selection of the 2D art presented in the museum and, when the artwork is recognized, it gives the visitor extra information. An example of that is shown in Figure 4. Besides art recognition, ArtLens 2.0 takes into account the personal preferences of the user and guides them to art that is in line with their preferences. The guiding is achieved with the help of mapping and beacon technology, leading the visitor through new pathways in the museum. The application is also modeled to have visitors engage with each other as they respond to the artwork.

During the development of ArtLens 2.0 a lot of the feedback from the first ArtLens app was taken into account. Furthermore, the second version uses the database from the museum for the art recognition. As soon as a new painting has been added to the database, the app is automatically updated as well. Overall the ArtLens 2.0 is a well developed app as it improves the visitor experience by guiding the entire visit, providing extra information and making the tour personal.

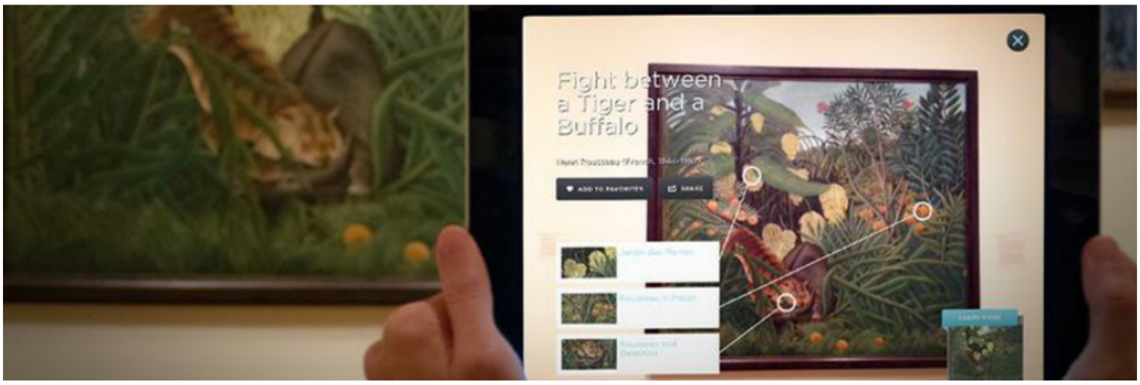


Figure 4: The ArtLens 2.0 annotates features on the artwork and provides extra information. Source: The Cleveland Museum of Art, 2013.

Skin and Bones

The second application discussed by Ding is the Skin and Bones application for the Smithsonian Museum of Natural History. This application aims to improve the Bone Hall exhibition, which has not been upgraded in over 50 years. The Bone Hall exhibition shows skeletons of many different animals. Before the Skin and Bones app, the information in the Bone Hall was given on the small signs near the skeletons. The text on those displays contained scientific terms that most visitors would not understand. The Skin and Bones app is designed to make that information more accessible and visible to the visitor and make the skeletons come to life. When the device is aimed at a skeleton, it is recognized and a 3D model is laid over it. The model visualizes how the animal looked like when it was alive and in some cases how it moved.

The visitors that use Skin and Bones stay remarkably longer in the Bone Hall, increasing the dwell time from 1:34 min to 14:00 min. Everyone who uses the app is very positive, but not many visitors use it as it is not clear to them that an app is available or that free WIFI is available to download the app. See Figure 5 for an example of the Skin and Bones app.

Blanton Museum of Art

The third and last application discussed by Ding is an app for The Blanton Museum of Art at the University of Texas in Austin. This university museum has an exhibition called “The Crusader Bible: A Gothic Masterpiece.” The pieces shown in this exhibition display only three languages: Persian, Judeo-Persian and Latin. These languages are hard to understand for English speaking visitors. To provide English translations, an AR application for an iPad is developed with the use of Layar, a free development platform

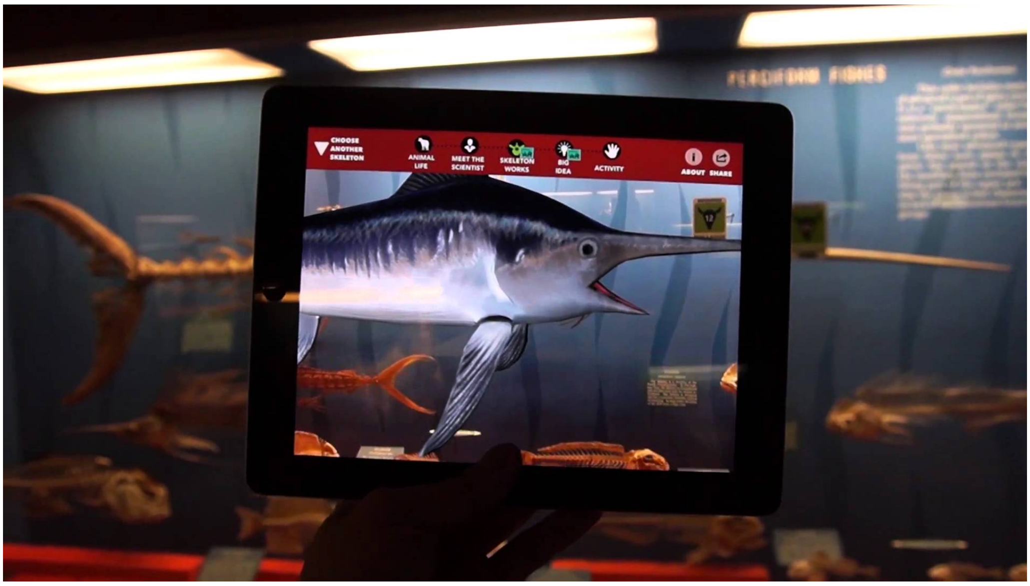


Figure 5: A model of the live animal is projected onto the seen skeleton. Source: Smithsonian National Museum of Natural History, 2015.

for AR browser applications. Only around 20% of the visitors use the application, because many of the visitors did not know what the iPad was for. Still, all of the users are very positive about the app and they stayed a little longer compared to the regular visitors.

Taipei Fine Arts Museum

Another example of mobile handheld AR for cultural heritage is given by Chang [7]. They do an experiment on user behaviour with visitors that use their art appreciation AR app during an exhibition in the Taipei Fine Arts Museum in 2012. The application runs on a 10 inch Android tablet PC. If the visitor points the device at a painting, the app recognizes it and gives a description of the painting. The description can be either audio or visual and it gives the visitor some background information on the painting. After the description part, the visitor can go into analysis. During the analysis the application allows the visitor to zoom in and out of the painting to go into more detail and get a more in-depth explanation. This teaches the visitor more about the painting. When the in-depth knowledge is gained, the app goes into the next step: the interpretation. The visitor can zoom in and out again and the app gives information on the general theme of the painting and how this theme is achieved. The last step is the judgment step in which the visitor can use his/her gained knowledge to assess similar paintings that are suggested by the app. Besides the suggestions, the app also provides some extra material and references of the painting the visitor is watching. Figure 6 shows the application with the user interface around it.

Figure 6: The seen artwork is shown on the tablet, with a user interface around it. Source: Image taken from [7].

During the exhibition 135 college students participated in the experiment. They were divided in three groups: an AR guided group, an audio guided group and a control group which had no guide at all. The experiment focused on the following dependent variables: learning, flow, time spent focusing on paintings, behavioral patterns from the visitor and the attitude regarding the use and acceptance of guidance systems from the visitor. The results showed that the AR guided group showed a significant increase in all of the dependent variables, except for the behavioral patterns and the acceptance of the guidance systems. The results in behavioral patterns from the AR guide group are similar to the audio guide group. Regarding the acceptance of the guidance system, the non-guided group would like to use a guide in the future, either audio or AR. The audio group would like to use audio again, but would also like to try the AR guide. The AR group would like to use AR again. Both guided groups also expressed that the guiding devices provide a good motivation to learn. The group without a guide expressed they missed one to help them understand the paintings.

Melaka heritage site

The last example is an AR application for a cultural heritage site in Melaka by Pendit [6]. The focus of their experiment is to create an enjoyable informal learning experience for the Melaka heritage site. The idea behind enjoyable informal learning is that the visitor gains knowledge, without feeling that they are learning. This is achieved by giving information in a way that is easy to process for the visitor. The application provides the user with an overlay of the world that annotates certain positions of the heritage site which can be clicked for more information. In an earlier paper by Pendit, guidelines of a conceptual model for enjoyable learning with AR for a cultural heritage site are given.

The development of the AR app was realized in three stages with the mentioned guidelines in mind. The first stage (pre-production) is used to design the user interface. During the second stage (production) the content is created. In the third stage (post-production) the user testing and evaluation are conducted. 200 visitors participate in the user testing, where they are asked to use the AR application and fill out a questionnaire afterwards. The questionnaire consists of 24 questions about informal learning, enjoyable informal learning and enjoyment. The questions are on a scale from 1 to 7, with 1 being the lowest and 7 the highest. On average the score was around 5.5, meaning that the visitors enjoyed using the AR application and learned something during their visit as well.

In the paper there is a mix-up in showing the results. From the three questionnaire sections, only two different tables are shown and one of those tables is duplicated. This makes it hard to deduce what kind of questions were asked and to replicate this study.

3.2 AR headsets for cultural heritage

Another way of showing AR is with a headset or an HMD. One of the better known headsets is the Microsoft Hololens. With a headset the user is able to use both hands and see the augmentations in the world, instead of on a screen, which creates a better immersion.

Explorative AR system Caggianese

The AR application by Caggianese is meant for exploring a city and find Points of Interest (POI) [8]. With the help of a see-through display the user can see the location of a POI in the world. The index finger of the user can be used as a pointer and when hovering for 2.5 seconds on a position, the system will recognize that as a click. This method is called the 'wait to click' method. If a user selects a location, he/she can start navigating to that position. When the user arrives at the location, he/she can get more information about the selected POI. Figure 7 shows how the selection of a POI looks.

The hardware used for this system consists of a see-through display combined with a 9DOF tracker, an RGB-D camera and a GPS sensor. All of the data these sensors acquire is sent to a mini PC. The 9DOF tracker is used to track head movements. Tilting your head backwards sets the system to idle or back to active. An idle system has all functionalities turned off, except for the tracker to turn it back on. When tilting the head forward the user activates the system control state. In that state the user can change parameters, such as what kind of POI to show. The RGB-D camera is used for the hand gestures. The depth images obtained from the depth camera are put in the tracking pipeline to get the tip of the index finger. The GPS sensor gets the position of the user to enable navigation.

The system is tested at a public smart city event in Napoli. There is no mention of how they conducted their experiment and what kind of questions are asked, only the lessons

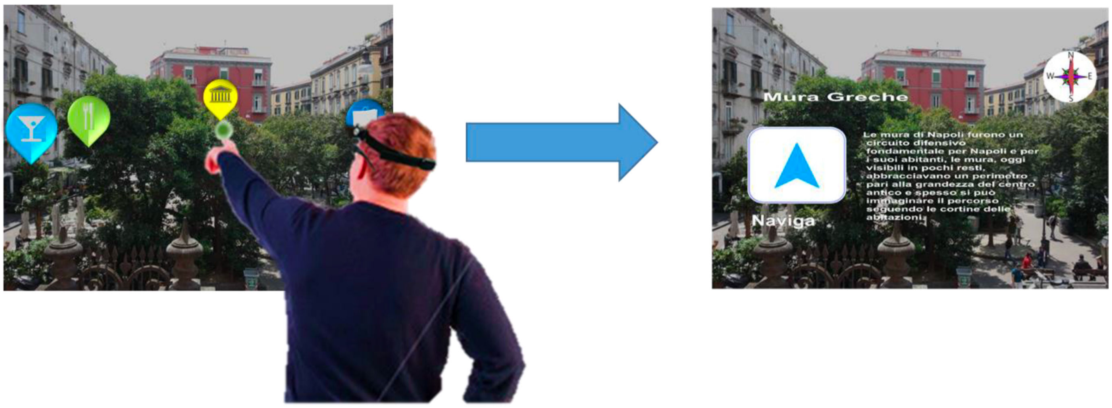


Figure 7: An infographic on the explorative AR system for Napoli. The user selects a POI and can start navigating. Source: Image taken from ICAR citing the paper from Caggianese [16].

learned are stated. One of those learned lessons is that the head tilting was not comfortable for many users. Another is that the augmentations are not visible with bright environment light. This was caused by the see-through glasses not blocking enough sunlight.

Explorative AR system Brancati

Brancati continued the development of the explorative POI system of Caggianese, but some of the hardware is different and the focus of the study has changed [3]. The focus of Brancati is the usability of the wearable AR system. For the hardware a META 1 is used, which has an integrated 3DOF tracker and an RGB-D camera. An external GPS sensor is used for the position of the user. The META 1 has no internal processing unit and is therefore connected to a Microsoft Surface tablet. There is no mention of how the tablet is integrated into the entire system or how it is 'worn' by the user.

The experiment was done during an exhibition in Naples. 35 participants are involved in the experiment and each participant gets a short introduction of the experiment and after that he/she is introduced to the system. First the participant was told about the different options in the user interface, followed by the different interactions that can be performed. After the introduction the participant is free to try out the system and ask questions about it. Once the participant felt comfortable in using the system, the experiment started. The participant has to select a POI at different distances and navigate to one of them. After the experiment the participant is asked to fill out some questionnaires.

The questionnaires used are the System Usability Scale (SUS), User Experience Questionnaire (UEQ), Usefulness Satisfaction and Ease of Use (USE) and the Device Assessment

Questionnaire (ISO 9241-9:2000 part 9 annex C). The SUS is a short usability questionnaire and the results of these are above average. The UEQ gives results in six different aspects of user experience. The different aspects are attractiveness, perspicuity, efficiency, dependability, stimulation, and novelty. All of these aspects scored very well, except for efficiency. The score of the USE questionnaire is above average as well. The Device Assessment questionnaire looks at the physical and mental effort required to use a system. Most of the items on this list scored above average, but some parts have a low score. This indicates that some parts of the AR system could be cumbersome to use. Participants enjoy using the system and the good system usability supports that.

Manchester Art Gallery

The last example of an AR headset for cultural heritage uses Google Glass in the Manchester Art Gallery, by Leue [9]. The experiment focuses on improving the learning experience from the visitors. The experiment is conducted in two stages. The first stage is to assess the user requirements and the second stage is focused at the learning experience. Because Google Glass covers only one eye, it is not a full AR experience, but it does use the HUD approach on the recognized objects. For a schematic of Google Glass, see Figure 8.

The 22 participants are recruited through social media from the Manchester Art Gallery. The experiment starts with an explanation of how Google Glass works, followed by viewing three paintings. The initial painting is chosen by the participant and from there he/she can choose from suggestions. The suggestions are based on the same artist, same medium or same theme from the first painting. When a suggestion is selected, navigation starts towards that painting. The application can recognize the painting and give more information about it. This information can be textual or can be read aloud through the bone conducting speaker. After viewing three paintings an interview is conducted.

The questions in the interview are based on the Generic Learning Outcome (GLO) framework. Many of the participants mention that they take a closer look at paintings. They take the time to look and listen to the given information and ask themselves questions, instead of just looking and walking away. Numerous participants also confirm that they see the added value of Google Glass in terms of learning more. Even though the learning experience is increased with Google Glass, there are some factors that limit the overall experience. The technology does not always work fluently as the sound volume was not high enough, the prism is small and difficult to adjust and some participants felt isolated whilst using Google Glass.

Despite the flaws, the overall conclusion is that with Google Glass, visitors learn more. However, this is all based on subjective information since there is no control group. Also, the experiment is in a partially controlled environment, where the participant has to follow a given set of paintings. This could have an influence on the learning experience.

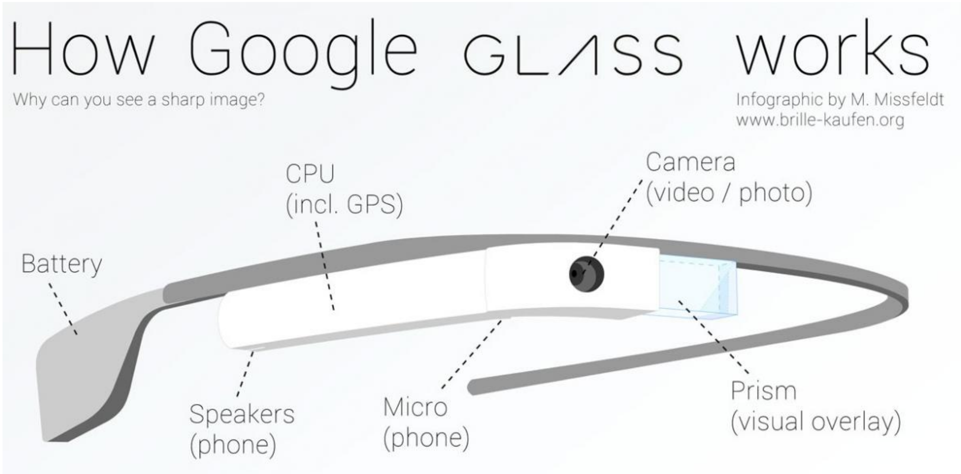


Figure 8: Short overview on how the Google Glass works. Source: Image taken from www.tech.de

3.3 AR in other domains

Important lessons can be learned from looking at AR in other domains besides cultural heritage. The following discussed papers are on those other domains.

Outdoor navigation

One of the many applications of AR is navigation, with an example from Kerr. They use a monocular HMD to navigate around an outdoor urban environment [4]. Their experiment aims to get a better understanding of the effect on user experience when using an AR HMD in an outdoor environment. As with the Google Glass, this system is not a full AR experience, as it only covers one eye for the AR and provides only a HUD.

The system consists of a headband with a see-through display attached to it that is in front of the left eye of the user. To the side of the headband, a smartphone is attached that provides information such as the orientation of the head, the location of the user and the graphics for the display. The user can interact with the system through hand gestures. These gestures are detected by a second smartphone that is worn on a lanyard around the neck of the user. Both smartphones are connected to each other via Bluetooth. The user interface consists of a top-down egocentric 2D rangefinder, with a target location annotated as a white dot. The target location is displayed as an icon in the field of view of the user. The user interface is seen in Figure 9.

Figure 9: A visual of user interface. Source: From the paper by Kerr [4]

A total of 8 people participated in the experiment. Each participant had to complete three tasks. The first task is navigating to a location. The location and distance is given and shown on the rangefinder and as an icon. The rangefinder and location icon are updated as the participant gets closer. Once arrived, the participant selects the acknowledgement icon and starts with the second task. The second task is navigating to another location, but before arriving, an update is given that changes the target location, which is the third task.

After the experiment the users fill out a questionnaire. Some important findings from this study are the external environment and situation awareness. When the participant is outside, external factors of the environment play a big role, such as sunlight. Sunlight is a major factor on the visibility of the projections on the display. With bright sunlight the projections are barely visible. Regarding the situation awareness, many participants noted that they had trouble with crossing a road. The monocular display limited their peripheral vision. They had to move their head a lot further than usual to get a good impression of the situation around them. In this study it becomes clear that the usage of outdoor AR with an HMD still needs a lot of improvement before it can be used.

Alien Contact!

Another domain for AR to be applied is education. Dunleavy conducted a study where AR aids education, seen from the perspective of both the teacher and students [11]. They developed a game named Alien Contact! that has a popular theme, namely aliens, taken from video games and the entertainment industry. The game is designed to teach math, language arts, and scientific literacy skills to middle and high school students. The game is played on a Dell Axim X51 handheld computer and with the GPS location, the real world position is correlated to the virtual position in the digital world. Whilst walking around the real world, the device shows a map with digital objects or persons as augmentation in the real world. If a student is close enough to a digital object or person, the game gives the player information. This is in the form of audio, video or text, which provide narrative, navigation and collaboration cues. The game also confronts the player with academic challenges. During the game, the students work in groups of four, each with their own role, and they have to solve why aliens have landed on Earth. By exploring the AR world, the students can find clues and have to work together to get all of the information and advance to the next stage. Alien Contact! is designed so that teachers can easily change some of the content in the future, such as the academic challenges or a different crisis.

The focus of the experiment is to see if AR can aid or hinder in the students their understanding of math and development of literacy skills. The focus from the teachers perspective is to see how they experience using AR for teaching. There is a total of 6 teachers and 80 students who participate in the experiment, spread over three schools in northeastern United States. All of the data is gathered through formal and informal interviews, direct observations, site documents and website postings. The interviews are conducted with all of the participating teachers and a subset of the participating students. From the results, it is found that students are engaged and motivated. This is mainly due to the fact that students got to work with new technology, going outside to play the game and the teamwork that was necessary to complete all the tasks. However, there are some problems with the hardware and logistics. There is also a cognitive overload for the students. Some problems with the hardware are that the GPS is not always accurate or working and sometimes the handheld would freeze. The logistics problems are that with only one teacher, handing out all the devices, helping teams with problems and keeping an eye out on the safety of everyone, is nearly impossible. The experiment is possible due to the fact that the researchers are there to help. One of those safety problems is that students are almost too engaged and don't pay attention to their physical surroundings. This can be dangerous in some areas around the school. The Alien Contact! game still needs improvement, but it can already be seen that it highly motivates students to play and learn at the same time.

SMART

Another example of AR applied in education is SMART, a SysteM of Augmented Reality for Teaching, by Freitas [12]. SMART is an educational system used in the 2$^{\text{nd}}$ grade on three schools to measure the impact on the learning capabilities of the students. The system consists of a PC with a monitor or data projectors to show an augmented video, a web camera for the detection of the markers and racquets that have markers attached to

Figure 10: Some of the 3D models that can be freely manipulated using SMART's racquets. Source: taken from the paper by Freitas [12]

both sides.

During the study several games are tested, but only two are explained. The first game is an animal classification game. In one hand, a racquet is held to see an animated animal and with the other hand another racquet is taken that corresponds to the animal's classification. The second game is also a categorization game, but with a classification of modes of transportation, such as aerial or terrestrial. If the answer is correct, the game provides audio feedback, in the form of an applause. In the case of a wrong answer, the game plays a buzzer sounds. An example of some models for the modes of transportation game is seen in Figure 10.

For the experiment, each class is divided in two groups: a control group that has a regular class and an experimental group that uses the SMART system. Before the experiment, there is a short introduction session with the teacher, to explain the system. The experiment itself is divided in three phases: the pre-test, the learning phase and the post-test. During the pre-test, the children get a short test about the subject they are going to learn about. The learning phase is the actual lesson, with traditional teaching methods for the control group and the use of the SMART system for the experimental group. After the lesson the post-test is performed, which is similar to the pre-test. The grade for the tests is given in percentages. To get the degree of learning (annotated by <g>), the following formula is used: <g>= (posttest - pretest) / (100% - pretest). The results show that the use of the SMART system improved the learning of many students. The only students that improved less with SMART than the traditional teaching methods, are the students that already score good. The results show that using the SMART system improves learning.

The improvement is not only because of the AR, but also due to the fact that it is a game.

3.4 Discussion

The reviewed papers on handheld AR for cultural heritage apply AR in different ways. All of them use a heads-up display (HUD), which is a digital overlay on the real world on the screen of the device. The overlay could be text boxes (as with Artlens 2.0) or like a user interface (as with the application for the Taipei Fine Arts Museum). A HUD is mainly used as a source of extra information on the object the user is viewing. It is also used for guiding the visitor, like with ArtLens 2.0 and the AR application for the Melaka heritage site. The Melaka application shows locations on top of the real world, indicating in which direction certain locations are. Besides the HUD, Skin and Bones also shows 3D objects placed on top of the viewed skeletons, that are normally hard to exhibit with live objects. Those objects are given as extra visual information to the user.

Looking at the AR headsets for cultural heritage, all reviewed papers use a HUD and are capable of guiding the visitor. In the case of the explorative system, the HUD shows the locations of interesting POI's and can give extra information on them. The paper on Google Glass for the Manchester Art Gallery does not explain how the navigation works.

In general there are two types of AR used in the reviewed papers, namely a HUD and 3D objects placed in the world. The HUD is used to provide extra information on an object that the user sees or as an indication of a location to guide the user. The 3D objects are used as extra visual information and since it is digital, it is possible to show objects that can otherwise not be shown. With regard to the development of the AR application for the RSH, the interior of the RSH does not offer enough interesting individual objects to use a HUD to give extra information and the house is small enough that navigation is not needed. However, there are many objects related to the RSH that are unavailable, but can be shown with AR. This makes the addition of digital 3D objects a good option.

Looking at the papers on AR in other domains, there are some lessons to be learned. The paper on the outdoor navigation system notes that the use of an HMD can reduce the peripheral vision of the user and with that the awareness of his/her surroundings. The paper on Alien Contact! also mentioned that users lost awareness of their surroundings, but that was due to the immersion into the application. In the case of the RSH, this is important. Everything inside the house is fragile and should not be touched. Observing users during the experiment could give an insight on how aware they are. The papers of Alien Contact! and SMART showed that AR motivates users and can help them learn more. This is a confirmation that AR for cultural heritage has a positive impact on the visitors.

So far all of the reviewed papers are focused on one device and evaluating that. There are no papers found that compare two different AR devices. It is also mentioned by Kerr [4] that a comparison between two AR devices would be interesting to see which is more appropriate in certain situations. This research is comparing two devices, giving a new insight on the area of AR.

Concluding the literature review, all of the reviewed papers on AR for cultural heritage show that the addition of AR to cultural heritage has a positive impact on the visitors. Visitors learn more from information that is overlaid on the real world, they take more time to look at art, they know how to navigate to certain locations, they know where a certain POI is and they are more motivated to learn something new. The reviewed papers on AR in other domains confirm that AR is helpful in certain situations. In the case of navigation it helps users get to a certain position with ease and in the case of education, students are motivated and learn more. However, those studies also show some downsides to the use of AR. It is possible that the technology can fail and whilst using AR, users tend to lose awareness of their surroundings. This could be either by a loss of peripheral vision with an HMD or being so immersed that the user only focuses on the device.

4 The Rietveld Schröder House

The experiment is held at the RSH in Utrecht and implemented in the current tour. This section gives a brief history of the RSH, Gerrit Rietveld and Truus Schröder, and explain how the current tour works and how AR could be an addition to that.

4.1 History

The RSH is built in 1924 by Gerrit Rietveld (1888 - 1964, the architect of the house) as an assignment from Truus Schröder (1889 - 1985, the patron and resident of the house) in Utrecht. Rietveld begins as a furniture maker and around 1918 he starts making furniture that is influenced by De Stijl, including the Red and Blue Chair as his most famous work. The RSH is his first work as an architect to design a complete new house. Schröder met Rietveld for the first time in 1921, when he remodeled a room in her house. After her husband passed away, she asked Rietveld to design a new house for her. Rietveld has his architecture firm inside the RSH from 1924 to 1933 and Schröder helps with the work of Rietveld. Schröder lives her entire life in the RSH and after she passed away in 1985, the house was renovated completely to look like it did in 1924 again and it is opened up to the public. In 2000 the RSH is added to the UNESCO World Heritage list.

What makes this house so special is that it is very modern for the time it was built. There are many small smart things inside the house that could help make living easier, because Schröder wanted to live in the house and feel free. To achieve that even more, and come loose from the Earth, the living part of the house was on the first floor, such as the living room and bedrooms. It is possible on the first floor to have it completely opened up, or to be closed in several compartments. This is achieved by moving the modular walls. With the ideas of De Stijl, Rietveld designed and painted the entire house, making it even more special. According to UNESCO there are two criteria why they placed the RSH on the World Heritage list: "... icon of the Modern Movement in architecture and an outstanding expression of human creative genius in its purity of ideas and concepts as developed by the De Stijl movement" and "With its radical approach to design and the use of space, the Rietveld Schröder House occupies a seminal position in the development of architecture in the modern age." [15]

4.2 Current tour

To visit and join a tour at the RSH, tickets need to be ordered online. Each day has five tours, except for Friday, which has nine tours and there are no tours on Monday. Each tour consists of a maximum of 12 persons and takes about an hour. Before the tour starts, all visitors are gathered in the ticket office next to the RSH and they get an audio guide and a floor plan of the house. Most of the time those tools are sufficient for visitors to roam around by themselves. At certain times, the tour guide tells something about the house and guides the visitors from the ground floor to the first floor. On the first floor the modular walls are demonstrated. Only the tour guide is allowed to touch anything inside the house, unless stated otherwise. Visitors are always allowed to ask the tour guide questions to get more information. When the time is up, the tour guide announces this and all the visitors go back to the ticket office and hand in their audio guide and floor plan.

So far the tour is limited to the objects that are inside of the house or that are available at all. There are many things related to Rietveld and Schröder that are gone. With the help of AR these things can be made visual again. Visitors are not allowed to touch anything inside the house, unless stated otherwise, but with AR interactive objects can be made. These interactive objects can be 'touched', giving visitors more freedom. An addition that can be made is more information on Truus Schröder. Currently the tour does not offer a lot of information on her and adding that as AR provides something new to the visitors.

5 Implementation of the two AR prototypes

This section covers the implemented AR prototypes for the experiment. These implementations need to be similar, to reduce their affect on the visitor experience. Therefore requirements are given that both implementations need to meet. A description of how they work and what they show during the experiment is given. To get a complete view of the work and choices, all considered options and limitations are given of the entire development.

5.1 Requirements

For the experiment two AR prototypes are implemented, one for the Galaxy S7 and one for the Aryzon. To reduce the affect of the application, both implementations need to be similar. the only difference is the device and how the device shows AR. The following requirements are set:

- The implementation is capable of showing 3D objects that are anchored to a position in the real world. The addition of AR to the RSH is to show new objects in the house and not providing a layer of extra information on the existing content.
- Only one marker image design is used in the RSH. That way the user needs to learn only one design as an indication of AR.
- The application logs data during the experiment on the device usage. This data is used to get an insight on the usability of the device.
- It is not possible to interact with the application or 3D objects. The experiment is aimed at the influence of adding 3D objects to the RSH on the visitor experience. Interacting with an application is very different between the two devices, possibly affecting the results of the visitor experience. Leaving out interaction limits the usage of the application and allows the visitor to focus more on the AR.

5.2 Implementation

The experiment is aimed at the comparison of the two different AR devices. They are compared on the visitor experience of two different participant groups, who each use a different AR device. See Figure 11 and 12 for an image of the Aryzon and the Galaxy S7 that are used during the experiment.

The two biggest differences between the two devices are the way AR is shown, as projections in the real world with the Aryzon and as a video on the mobile handheld, and the field of view (FOV). In this case, FOV refers to the AR part of the view and how much the user can see of that. For the Galaxy S7 this is limited by the camera of the phone, as only in that view AR is added on the screen. The FOV of the Aryzon is smaller compared to the Galaxy S7, as the projections are stereoscopic. To achieve the stereoscopic view, the phone display is split in two, creating images for both eyes. This halves the width of the AR view, decreasing the field of view. See Figure 13 and 14 to see the FOV of both devices at a similar distance. See Figure 15 to see through the Aryzon.

Both implementations are build with Unity 2017.2 and Vuforia that is integrated with it. Vuforia is an AR SDK and works by detecting markers and adding an AR object to that, anchored to that position in the real world, meeting one of the requirements. It is possible to detect images, cuboids, cylinders and 3D objects. In the case of the RSH, the markers should not be too intrusive and disrupt the experience of the RSH, therefore images are

Figure 11: The Aryzon AR cardboard.

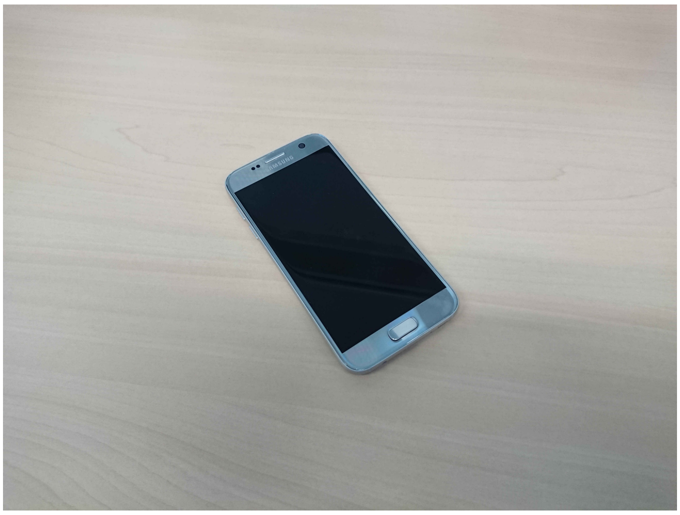


Figure 12: The Samsung Galaxy S7.

Figure 13: The view of AR on the Galaxy S7. The entire magazine and the note are visible.

Figure 14: The view of AR on the phone for the Aryzon. Both sides of the display are combined for the stereoscopic view. Almost the entire magazine is visible, but the note is partially visible.

Figure 15: Seeing through the Aryzon.

chosen as a marker. These are easily placed and can be designed to fit in the atmosphere of the RSH. Vuforia has two types of markers, one being custom images and the other being VuMarks. The custom images can be designed as anything, but one marker design can only show one object. A VuMark is like a QR code embedded in a logo. One VuMark design can contain many different QR codes and allows one logo design to show many different objects. See Figure 16 and 17 for the custom marker and the VuMark that are designed for the experiment.[1]

The development for the Aryzon works with Vuforia as well and has the addition of the Aryzon SDK, which changes the view to a stereoscopic view on the phone with a black background. This is necessary for it to work with the Aryzon cardboard. See Figure 14 to see the stereoscopic view.

The experiment is held in two different rooms in the RSH, both showing a different object. The requirement is that only one marker image design is used in the RSH, which is the marker seen in Figure 16. This marker is choses as it follows the style of the RSH and the Centraal Museum. A custom marker image is capable of showing only one object, therefore a secondary source of information is needed to determine the location of the user. To get the location of the user, beacons can be used. A beacon sends a signal to a receiving device. With that signal, the device has information on that beacon. Part of that information is the distance to the beacon in centimeters. To get an approximation of the location of the user, beacons are placed in each of the rooms. The closest beacon to the user determines the room he/she is in. With the use of beacons, one custom marker image can be used and different objects are shown depending on which room the user is in. However, testing showed that the distance between the beacon and the device is

[1] These designs are made by Joost Overbeek.

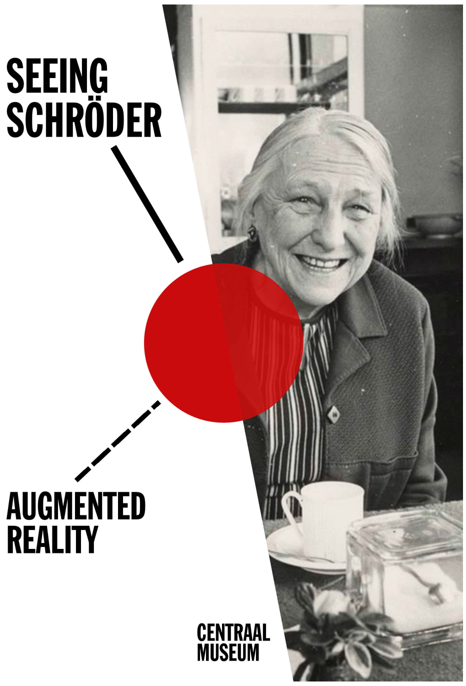


Figure 16: The custom designed image marker.

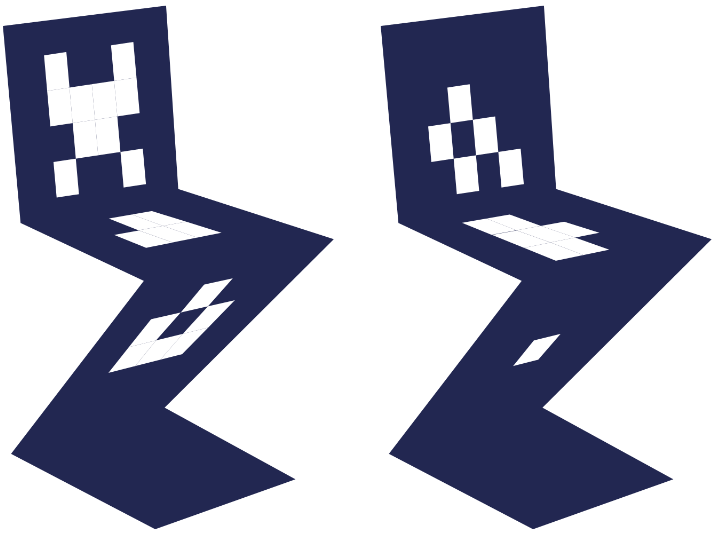


Figure 17: Two VuMarks that have the same design, but each with a unique QR code embedded in it.

Figure 18: These shoes are placed in the hallway during the experiment.

updated slowly, making it hard to use in a small environment. Due to that reason, the use of beacons is omitted. Without a secondary source of information, custom image markers can not be used. Instead, a VuMark is used, as it consist of one design that can have multiple QR codes embedded in it. The design of a VuMark is limited, making it harder to create one that fits the style and atmosphere of the RSH. The design of the VuMark is based on the ZigZag chair, a famous chair design by Rietveld, making it a fitting addition to the RSH. The VuMark is seen in Figure 17.

In the background, data is collected on the usage of the application. The application logs certain actions the user performs with a timestamp. These actions are starting the experiment, detecting a marker image, losing track of a marker image and ending the experiment. All that data is written to a file for later processing.

To the user, the applications are capable of detecting two different markers, each showing a different object. The objects are giving more information on Truus Schröder, as both tell a small fact about her. The first marker is placed in the workshop and shows a magazine that Truus Schröder worked on, seen in Figure 13. The other marker, placed in the hallway, shows shoes, seen in Figure 18. Both objects have a digital note next to them as a source of information to the user. The current audio tour does not cover this information, making the AR an addition to the tour. By pointing the camera of the device, either with the Galaxy S7 or the Aryzon, at a marker, the image is detected and an object is placed on top. The object is anchored to the position of the marker, allowing the user to walk around it and see it from up close. There are no other functionalities in the application, meaning there is no interaction possible with the object or any other form of interaction with the application itself.

5.3 Considered options and limitations

During the development of the prototypes, several options are considered and limitations are found. These options are previous ideas and insights that did not make the final implementation, but are valuable lessons on what works and what does not.

Early versions of the implementation are run on a Sony Xperia M5 handheld device. However, it is not possible to build an application on the Xperia M5 that is build with Unity 2017. A new device, a Samsung Galaxy S7, is chosen, as it is capable of running applications build by Unity 2017 and runs AR applications smoothly. In addition, a Galaxy S7 is capable of running the object scanner from Vuforia, which is necessary to implement object detection.

With the recent release of the Aryzon, the SDK is not optimized yet. This is seen in the placement of a virtual object on an image marker. The object is not placed on the position of the marker and has an offset. Increasing the accuracy of the object placement required time, with results that are still not perfect.

Looking at AR with an Aryzon, a small dot is seen in the center of the users vision and moves along with him/her. This dot is used to interact with objects seen in AR. There are two possible actions that a user can have with an object. The first action is when the dot hovers on an object. This could set some action in motion, such as making an object rotate. The second action is when an object is 'clicked'. Clicking an object is done by the 'wait-to-click' method. The user needs to hover the dot for a certain amount of time over an object, usually between two to three seconds, and after that the object is clicked. That could also set some action in motion, such as showing extra information of the object that is clicked. Interacting with an object on a regular smartphone is done by tapping the object on the screen. This action is more intuitive to users than using a gaze dot. To take out the affect of these interactions on the outcome of the experiment, it is decided that the final implementations have no interaction.

One of the functionalities Vuforia offers is extended tracking. If a user sees a marker, but moves away and loses sight of the marker, the object can still be seen. By tracking the background, Vuforia predicts the position of the object inside the world. This feature is useful for larger objects. However, extended tracking does not work accurate. With slow device movement, Vuforia is capable to do the extended tracking correctly. With sudden changes or a moving background this is not possible and the objects are shown incorrect, meaning that objects are rotated or misplaced. If the marker is detected again, the object is placed back on its original position. Since extended tracking is not stable, it is not used in the final implementations and the objects shown in the experiment are small enough that they are seen completely with the marker in sight.

Looking at the future of an AR tour in the RSH, the option of choosing an AR tour is implemented. This allows the user to have a more personal experience. By looking at a certain marker at the start, a tour is set and that determines which objects are shown on the markers. If a different tour is chosen, those same markers show different object. For the experiment, the option of choosing a tour was not necessary and therefore omitted from the final implementations.

With Vuforia it is possible to detect several markers at the same time, each showing their corresponding object. The condition for this is that all markers are unique and can be distinguished from each other. This is easy to achieve with custom designed markers, but with a VuMark this is not possible. VuMarks all have a unique QR code, but are recognised by Vuforia as the same image marker. By detecting two different QR code VuMarks, Vuforia wants to show the corresponding objects on the VuMark image. This results in showing the objects that correspond to both ID's on both VuMark image markers. This problem is a drawback of using VuMarks.

To reduce the intrusiveness of image markers, one of the options is to use object detection. Enabling the detection of already existing objects in the RSH would take out the need of image markers. The object scanner from Vuforia scans 3D objects that need to be detected. However, the object scanner is limited to small models that are around the size of an A4 paper in width and length. Using scaled models could solve this problem. Testing the object detection shows that within the object scanner the detection works. However, if the file is exported to Unity, the application is not able to detect the object. The reasons for that are unknown and object detection was omitted from the final implementations.

Implementing the use of beacons is done by using the iBeacons asset from the Unity asset store. This allows a device to receive signals from sending beacons and get the information out of it. This information contains the distance to the beacon in centimeters, the received signal strength indicator (RSSI) and other values that are not relevant. As mentioned before, the distance to a beacon is updated slowly and is therefore not accurate and fast enough to be used in a small environment. The RSSI value can be used as an indication of distance as well, but this is only accurate up to one or two meters. Beyond that distance, the RSSI value from all beacons are in the same range and the application cannot determine a closest beacon. Due to time it is not possible to find a solution for this, but some directions to solve it are given:

- To improve the distance value, linear prediction can be used. If a distance value changes in a few intervals to one direction, the next values are predicted and are used as an indication of the distance to the beacon. This gives a faster result on the possible distance of the user.
- To improve the RSSI value, the value that indicates where the user is closest to, is kept as closest value, until another beacon shows values that indicate the user is getting closer to it. These threshold values are important and need to be found first.
- The last idea is to combine both values. The RSSI quickly shows to which beacon the user is the closest and that beacon is set as the closest. After a few updates of the distance, the distance value is used as indicator of which beacon is closest, taking out the unstable values of the RSSI.

6 Methodology

Related work shows that AR has a positive impact on the visitor experience, therefore a comparative study between the two AR devices is conducted. There won't be a control group that doesn't use an AR device. The participants are observed to see how they use the device and how well they are aware of their surroundings. These observations are to see how the user handles a device and everything inside the RSH is fragile and all objects are museum pieces and should not be touched.

6.1 Prerequisites

This experiment looks into the visitors "enjoyment", "satisfaction about their visit" and "knowledge gained", subsequently referred to as *satisfaction*, *enjoyment* and *knowledge*. These dependent variables influence the independent variable: the AR device. It is important that the dependent variables are measured for one device per visit, therefore a between-subject design is chosen. The experiment is conducted at the RSH during a tour, meaning there is a high external validity. Only one AR device is available during the experiment, thus only one visitor at a time can use it. More visitors can use it, by participating after each other. It is possible that a visitor sees how the application is used before participating, gaining prior knowledge and introducing noise. However, this would also be the case if it is implemented in the tour outside of an experiment, thus the prior knowledge only increases the external validity. To produce results that give a good indication of the device preference, a minimum number of participants is set to 15 for each device. The experiment is held during two days, with one day to test the Aryzon and the other day to test the Galaxy S7. This way both devices experience the varied audience of the RSH of an entire day.

6.2 Participants

Before the start of a tour, the visitors are gathered in the ticket office. An explanation is given by the tour guide on how the tour works. After that, the visitors are asked if they want to participate in an experiment to try out a new AR application for the RSH. The experiment starts in the workshop, one of the rooms on the ground floor of the RSH, and everyone that passes by can participate.

Even though everyone is allowed to participate, there are some limitations. The experiment is conducted in either English or Dutch, thus only the visitors that are capable of speaking English or Dutch, and reading English can participate. The RSH has visitors from all over the world and not all of them speak English or Dutch, reducing the amount of people that can participate each tour.

6.3 The experiment

Two rooms in the RSH have AR added to them. These are the workshop, where visitors can start, and in the hallway. The marker in the workshop is placed on the table with the "De Werkende Vrouw" magazine and the marker in the hallway is placed on the ground with the pair of shoes. It is placed on the ground to get an insight on how participants behave with markers that are placed on harder to reach locations. Before the experiment starts, a short introduction is given to the participant on how the AR application works. During the experiment the visitor is asked to think out loud and tell what he/she is able to see and how they feel about it. A short explanation is given on what the participant sees and why that object was chosen. The same is asked of the participant with the second AR object in the hallway. After the experiment, the AR device is taken back and a next visitor can start the experiment.

To allow visitors more freedom with the Aryzon, there are no headbands used to wear the device. Instead, users can hold it with their hands and keep it in front of their eyes in the same position as if wearing it. This also reduced some of the time during the experiment to get the Aryzon in a comfortable position on the users´ head. However, this does take away some of the information on how comfortable the Aryzon is and how it would affect the awareness of the user.

6.4 Data gathering

To gather all the data, three methods are used: a questionnaire, observations and data logging on the device. The questionnaire is filled out after the tour when the visitors are back at the ticket office, the observations are made by the experimenter during the experiment and the data logging is done by the device. Since the next tour starts soon after a tour has ended, there is no room for an interview with the participants.

Questionnaire

The questionnaire collects the data on the dependent variables and the demographic information of the participant. It is designed with five sections, each with its own topic. Each of the following sections are explained below. The full questionnaire is found in Appendix B.

The demographic information section includes questions about the age and gender of the participant. To procure some background information on the participant, it also includes question on the reason for the visit to the RSH, if he/she visited the RSH before and if there is previous experience with technology such as AR.

For the enjoyment section, the questionnaire by Lin, Gregor and Ewing is adopted[10]. They give a general definition of enjoyment, which consists of engagement, positive affect and fulfillment, and design a questionnaire to measure that. We adopt that questionnaire with some slight changes. The variable to measure on is the RSH and all three sections have three questions instead of four, to shorten the time required to fill out the questionnaire. All questions are on a 7-point Likert scale (1 = strongly disagree, 7 = strongly agree).

For the section about the satisfaction, the research presented by Jeong and Lee is taken into account [5]. Their research looks at how the museum environment influences the visitors' satisfaction about their visit. The visitors self-reports their satisfaction about the different parts in the museum environment, such as the content of the exhibition and the method of exhibition. The AR implementations have an influence on these variables, by showing new content and showing it in AR. Therefore these questions are adopted into the questionnaire. Beside the aforementioned variables, we also add a question on the entire visit to the RSH. All questions use a 7-point Likert scale (1 = very unsatisfactory, 7 = very satisfactory).

The section about the gained knowledge is the last of the dependent variables and for that the questionnaire by Pendit, Zaibon and Abubakar is used as inspiration [6]. They measure the informal learning of a visitor that visits a cultural heritage site. The questions for this questionnaire are designed with that in mind. The first question is if the participant was motivated to learn about the RSH. The second question is about whether or not the AR application helped the participant learn more. The last question asks if the participant learned something from his visit to the RSH. These questions are answered on a 7-point Likert scale (1 = strongly disagree, 7 = strongly agree).

The last section concludes the questionnaire. The visitor grades his/her visit to the RSH, can give feedback and comments on the AR device and can give additional comments about the entire experiment. It also includes two consent remarks that the participant needs to check, before the data of the questionnaire can be used.

Observations

The participant is asked to think out loud and say what he/she sees and thinks, to get feedback on how the participants experiences the AR. There are also notes made on how well aware the user is of his/her surroundings and how the device usage is. This information is used to get an insight of how users react to AR, how they handle the device and how aware they are. The device usage is an indication of the usability.

Data logging

As mentioned above, during the experiment the applications log data of actions the user makes. These actions consist of when the experiment starts, when and how long the user sees an AR object and when the experiment stops. This information is used to get an indication of how easy it is to point the device at the marker to see the AR object. If the experiment took a long time, including a lot of time looking at AR objects, it can be an indication that the user is more interested.

7 Results and Discussion

In this section all the gathered data from the questionnaires, data logging and observations, is evaluated and the AR devices are compared based on the visitor experience. The comparison of the visitor experience is split into the three factors: enjoyment, satisfaction and knowledge. Each factor is evaluated with SPSS Statistics 24 and a preferred device through statistical means is given. The logging data, observations and the final grade from the questionnaire are evaluated and combined with the statistical results.

7.1 Evaluation plan

The experiments have a total of 28 participants, where 13 use the Arzyon and 15 use the Galaxy S7. Both experiment groups have a good distribution in male and female. The Aryzon group has 6 male and 7 female participants and the Galaxy S7 group has 7 male and 8 female participants. The distribution of the age varies between the two groups. See Table 1 for the age distribution. The reason for this is that during the second day, 3 of the 5 tours are visited by groups of a high school students as part of an excursion. Due to the time, no more experiments are conducted and the difference in age distribution is neglected. Only three participants never heard of AR before, all others at least knew of the technology. This reduced the time required to explain participants what AR is and what exactly is happening when using the application. Two of the participants have visited the RSH once before and all other participants have no prior knowledge of the tour at the RSH, meaning there is no bias on the visitor experience from previous visits. Looking at the reasons to visit the RSH, besides the high school groups that have 'excursion' as a reason, almost all others have recreational reasons. These include an interest in the architecture, the Stijl, or as a day out. A few participants are visiting for study and research reasons. All together, almost every participant is interested in the house and everything related to it.

Looking at the totality of the demographic information, both experiment groups are close to equal, if we neglect the difference in age. Both groups can now be safely compared on the visitor experience, without taking into account demographic information.

Age category	0-17	18-24	25-39	40-59	60+
Aryzon	2	1	5	3	2
Galaxy S7	8	2	0	2	3

Table 1: Age categories distributed over the Aryzon and the Galaxy S7.

The comparison for each factor of the visitor experience is done on each individual question. This creates sample sets of size 13 or 15 for each question, with 13 for the Aryzon and 15 for the Galaxy S7. If a device is preferred by a majority of the questions in a factor, that device wins on that factor. The AR device that is preferred by at least two out of the three factors is determined to be preferred by the visitors. Instead of comparing on each question, all questions can be grouped together to compare on the entire factor. An advantage of that is an increase in the sample size, which can have a positive influence on the statistical test. This creates a sample set size of 117 on the Enjoyment factor for the Aryzon and 135 for the Galaxy S7. Satisfaction and Knowledge have 39 and 45 for the Aryzon and Galaxy S7 respectively. However, a disadvantage is that a high score on a few questions can increase the mean of an AR device sample set higher than the mean of the other AR device sample set, even if it is preferred by less than halve of the questions. Therefore comparing on the individual questions is chosen over comparing on the entire factor. Two statistical tests are used for the comparison: the Students t-test and the Mann-Whitney U test.

7.2 Evaluation with the t-test

Before doing the t-test, the sample sets are checked for outliers using boxplots. SPSS creates boxplots with regular outliers, shown as a circle, and extreme outliers, shown as an asterisk. One extreme outlier is found in the Enjoyment factor on the question Pleased for the Galaxy S7. See Figure 19 for the boxplot. The outlier is removed to enable the use of the t-test.

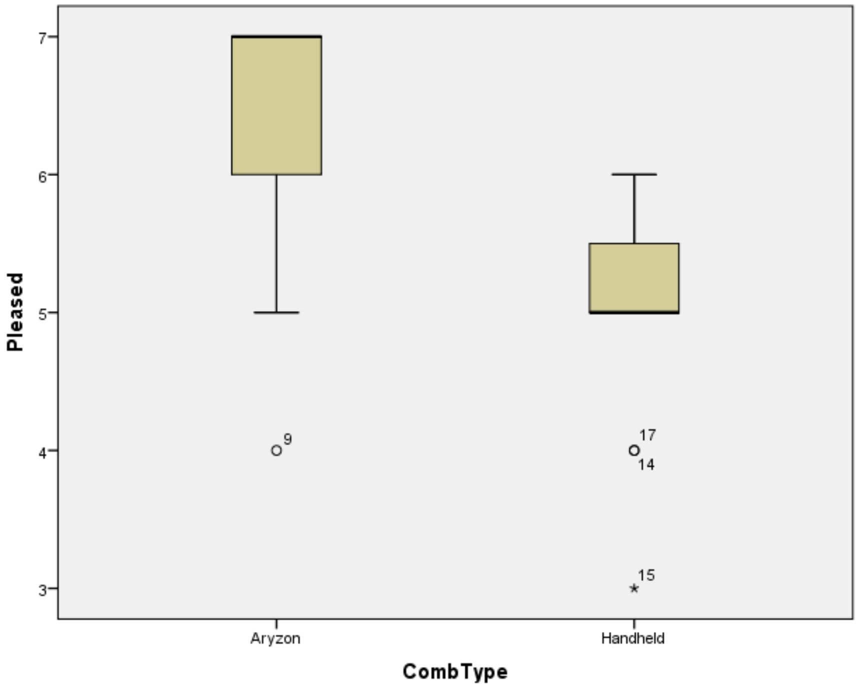


Figure 19: Boxplot of the question 'Pleased' with the Aryzon on the left and the mobile handheld on the right.

Using a t-test on a small sample size requires that it follows a normal distribution. As all sample sets have a maximum of 15 elements, the Shapiro-Wilk test is used to check the normality. The results are seen in Table 2. An important note is that the degrees of freedom (df) for the Pleased Handheld is 14, because of the extreme outlier discussed earlier. A sample set follows a normal distribution if the Sig. value is larger than 0,05. Those values are presented in bold. Only a few of the sample sets follow the normal distribution, meaning that it is not possible to perform a t-test on these samples.

	AR Device	Statistic	df	Sig.
Immersed	Aryzon	0,894	13	**0,110**
	Galaxy S7	0,805	15	0,004
Focused	Aryzon	0,875	13	**0,062**
	Galaxy S7	0,862	15	0,026
Concentrated	Aryzon	0,847	13	0,026
	Galaxy S7	0,927	15	**0,246**
Happy	Aryzon	0,776	13	0,004
	Galaxy S7	0,914	15	**0,156**
Pleased	Aryzon	0,776	13	0,004
	Galaxy S7	0,801	14	0,005
Contented	Aryzon	0,744	13	0,002
	Galaxy S7	0,866	15	0,030
Fulfilling	Aryzon	0,776	13	0,004
	Galaxy S7	0,926	15	**0,235**
Useful	Aryzon	0,870	13	**0,052**
	Galaxy S7	0,931	15	**0,279**
Worthwhile	Aryzon	0,619	13	0,000
	Galaxy S7	0,865	15	0,028
Content	Aryzon	0,803	13	0,007
	Galaxy S7	0,816	15	0,006
ShownContent	Aryzon	0,772	13	0,003
	Galaxy S7	0,882	15	**0,050**
Visit	Aryzon	0,638	13	0,000
	Galaxy S7	0,881	15	0,049
Motivated	Aryzon	0,816	13	0,011
	Galaxy S7	0,874	15	0,039
ARhelped	Aryzon	0,905	13	**0,158**
	Galaxy S7	0,918	15	**0,179**
Educational	Aryzon	0,809	13	0,009
	Galaxy S7	0,667	15	0,000

Table 2: Normality check results from SPSS. The sample sets that follow a normal distribution have their Sig. value presented in bold.

As a rule of thumb, sample sizes of 30 or larger don't require a normality check. To increase our sample size, the comparison is done on the entire factor.

A last check before the t-test is to see if there is a homogeneity of variances. This is tested with the Levene's test. If the p-value of the Levene's test is larger than 0,05, there is a homogeneity of variances between two sample sets. Table 3, shows the Levene's test F and the corresponding p-value of Levene's test for all three factors. All three factors have a homogeneity of variances, therefore the t-test can be performed.

Levene's test	Enjoyment	Satisfaction	Knowledge Gained
F	1,92	0,166	0,492
p-value	0,167	0,685	0,485

Table 3: The Levene's test on all three factors

Inputting the data into SPSS returns the results shown in Table 4. If the p-value is smaller than 0,05 there is a significant difference, showing that one device is preferred over the other. All three factors are smaller than 0,05 and the Aryzon has the higher mean value on all three factors, thus the Aryzon is preferred on all three factors.

Factor	AR device	N	Mean	Std. dev.	t	p-value (2-tailed)
Enjoyment	Aryzon	117	5,98	1,083	7,674	0,000
	Galaxy S7	135	4,96	1,025		
Satisfaction	Aryzon	39	6,28	0,887	3,954	0,000
	Galaxy S7	45	5,51	0,895		
Knowledge	Aryzon	39	5,87	1,218	3,090	0,003
	Galaxy S7	45	5,09	1,104		

Table 4: The t-test results of the three visitor experience factors.

7.3 Evaluation with the Mann-Whitney U test

The results obtained with the t-test are not ideal, as there is a preference on comparing the devices on each individual question. As a second test, the Mann-Whitney U test is chosen. It can be used on small sample sizes, as long as the two sample sets that are compared have a similar distribution shape. If the sets have a similar shape, the Mann-Whitney U test compares medians, otherwise it compares on the mean ranks.

To see if two sample sets have a similar shape, the boxplots and the histograms of those sets should have a similar shape. For example, in Figure 20 the histograms of both devices and boxplot of the Immersed value are shown. It is clear that the histograms differ in shape and this is confirmed by the boxplot. In Figure 21 the histograms and boxplot of the Focused value are shown. Here the histograms look very similar. The difference is that the Aryzon histogram has values from 2 to 6 and the handheld histogram has values from 3 to 7. This is seen in the boxplot as the Aryzon boxplot is shown higher than the handheld boxplot.

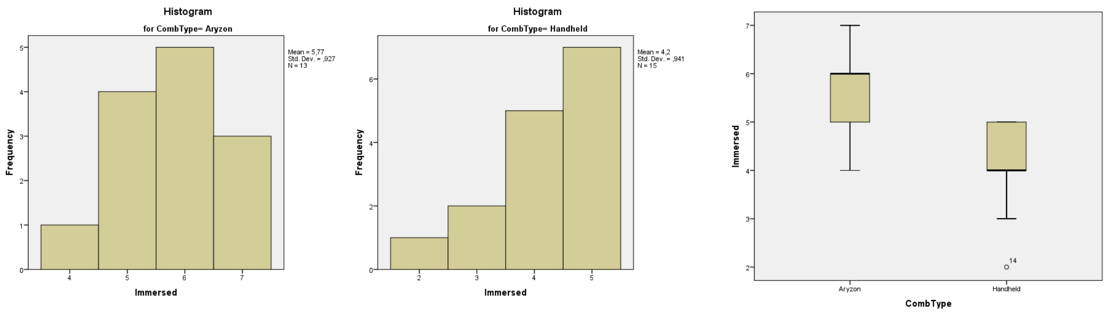


Figure 20: The left image shows the histogram of the Immersed value for the Aryzon. The center image shows the histogram of the Immersed value for the handheld. The right image shows the boxplot of the Immersed value for both the Aryzon (left) and the handheld (right).

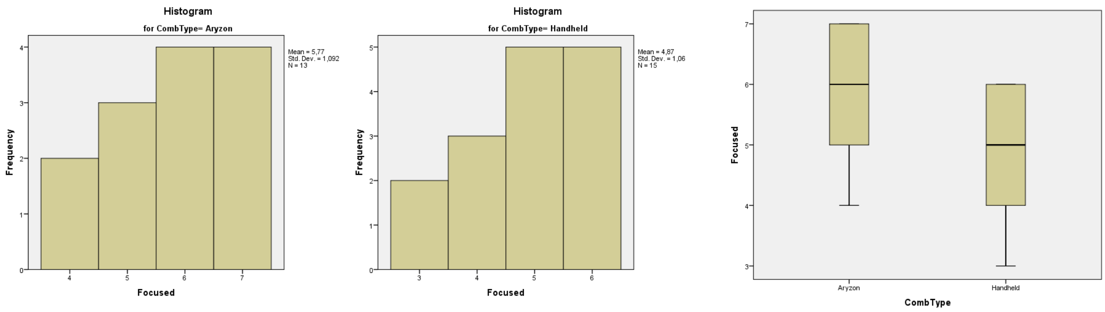


Figure 21: The left image shows the histogram of the Focused value for the Aryzon. The center image shows the histogram of the Focused value for the handheld. The right image shows the boxplot of the Focused value for both the Aryzon (left) and the handheld (right).

To see all the histograms and boxplots, see Appendix A. Viewing all the histograms and boxplots, it is clear that there are very few similar distribution, therefor the Mann-Whitney U test is performed by comparing the mean ranks.

With a small sample size, the exact method is used and not the asymptotic method. Since the sample sets have a lot of equal values, correction for ties is applied as well. The results are presented in Table 5. If the p-value is below 0,05, the difference between the two sample sets is significant and a device is preferred. These values are presented in bold in the table. Looking at the p-values, each factor has at least half of the values below 0,05 meaning that there is a preferred device on each factor. The preferred device is the Aryzon for all factors.

	AR device	N	Mean	Std. dev.	Mean Rank	Sum of ranks	MWU	p-value
Immersed	Aryzon Galaxy S7	13 15	5,77 4,20	0,927 0,941	20,19 9,57	262,50 143,50	23,500	**0,000**
Focused	Aryzon Galaxy S7	13 15	5,77 4,87	1,092 1,060	17,73 11,70	230,50 175,50	55,500	**0,049**
Concentrated	Aryzon Galaxy S7	13 15	5,23 5,13	1,166 1,060	15,12 13,97	196,50 209,50	89,500	0,724
Happy	Aryzon Galaxy S7	13 15	6,23 5,13	1,013 0,990	18,73 10,83	243,50 162,50	42,500	**0,007**
Pleased	Aryzon Galaxy S7	13 15	6,23 5,00	1,013 0,845	19,31 10,33	251,00 155,00	35,000	**0,002**
Contented	Aryzon Galaxy S7	13 15	6,31 5,27	0,855 0,961	18,73 10,83	243,50 162,50	42,500	**0,007**
Fulfilling	Aryzon Galaxy S7	13 15	6,23 4,67	1,013 1,113	19,62 10,07	255,00 151,00	31,000	**0,001**
Useful	Aryzon Galaxy S7	13 15	5,62 5,13	1,387 1,187	16,19 13,03	210,50 195,50	75,500	0,318
Worthwhile	Aryzon Galaxy S7	13 15	6,46 5,27	0,877 0,799	19,35 10,30	251,50 154,50	34,500	**0,002**
Content	Aryzon Galaxy S7	13 15	5,92 5,47	1,115 0,915	16,96 12,37	220,50 185,50	65,500	0,120
ShownContent	Aryzon Galaxy S7	13 15	6,38 5,40	0,650 0,910	18,96 10,63	246,50 159,50	39,500	**0,006**
Visit	Aryzon Galaxy S7	13 15	6,54 5,67	0,776 0,900	18,42 11,10	239,50 166,50	46,500	**0,011**
Motivated	Aryzon Galaxy S7	13 15	6,15 5,07	0,987 1,280	18,35 11,17	238,50 167,50	47,500	**0,016**
ARhelped	Aryzon Galaxy S7	13 15	5,23 4,80	1,589 1,265	16,08 13,13	209,00 197,00	77,000	0,348
Educational	Aryzon Galaxy S7	13 15	6,23 5,40	0,725 0,632	18,88 10,70	245,50 160,50	40,500	**0,005**

Table 5: Descriptives and results of the Mann-Whitney U test. The MWU is the statistical value of the Mann-Whitney U test. If the p-value is smaller than 0,05 the two sample sets are significantly different.

7.4 Data logging, observations and grade

The statistical results from the questionnaire conclude that the Aryzon is the preferred device. However, how is the awareness of the users, the usability and are there any technical issues? These are retrieved from the data logging and observations and discussed below.

The grade that is given is evaluated as well.

Data logging

The average time that a participant spent testing and looking at AR is 114,65 seconds and 73,83 seconds for the Aryzon. Participants of the Galaxy S7 spent 109,81 seconds testing and 83,44 seconds looking at AR. The ratio of AR to testing time is a lot higher for the Galaxy S7 (77%) compared to the Aryzon (64%). This could indicate that the participants using the handheld have an easier time detecting the marker images and that the Aryzon is harder to use. Another reason is that the handheld is a bit more intuitive, whereas the Aryzon required more explaining during the experiment, increasing the total testing time.

Looking at the amount of times that a certain AR object is seen, the magazine object varies a lot between the Aryzon and the handheld. For the Aryzon, the magazine is seen a total of 24 times and for the Galaxy S7 this number is 51. Part of the reason is that participants using the Galaxy S7, moved the device away from the marker more often to see what would happen. With a handheld this is easier than with the Aryzon, again indicating that the Galaxy S7 is easier to use and move around with. Looking at the average time spent looking at the magazine, this is 26,33 seconds for the Aryzon and 16,72 seconds for the handheld. The shorter time for the handheld confirms that participants moved more with the device compared to the Aryzon. The difference for the shoes object is too small to be taken into account.

Observations

The observations discussed here regard the things that the experimenter sees during the experiment and what comes up during the interaction between the experimenter and participant. First the observations about the Aryzon are discussed and after that the observations about the Galaxy S7.

A participant with glasses had trouble using the Aryzon, as it could not fit over its glasses. This made the use of the Aryzon rather annoying. Several times it was noted that the AR images and text are hard to see and read. The images would be hazy, indicating a wrong focus point. Sometimes the participant would see the image twice, with a slight offset between them. This indicates that the image from the mobile phone to the see-through glass is not correctly transformed. Some participants 'solved' this by holding the Aryzon a bit away from their eyes. To get a good view of the AR image, the participant needs to be relatively close to the marker image. This could lead to uncomfortable situations, depending on the placement of the marker image. This is mainly seen in the hallway with the marker on the floor. Users would crouch to be able to see the AR object. If there are more users with an Aryzon in the same room, they could block each other by being close to the marker image.

Whilst using the handheld, some participants had trouble reading the text on the small screen of the phone. They had to get closer to the marker image, but are able to do so in a rather comfortable way, since they could freely move the device and still be able to see the objects rather well. Others try to interact with the device by tapping the screen or trying to zoom in, which are interactions that are not implemented. This does show that using the handheld triggers interactions that are intuitive and expected. An interesting observation is that some participants noted they would like to use AR glasses, even though they didn't know about the Aryzon.

To sum this up, the Aryzon is harder to handle and has some technical flaws as the AR objects are sometimes hard to see. The Galaxy S7 is intuitive to use, easier to handle and has clear images of AR. There have been no indications of users not being aware of their surroundings or that using one of the devices has ergonomic issues. So in terms of usability of the device and the technological aspects, the Galaxy S7 is the better device.

Grade

In the last section of the questionnaire, participants are asked to grade their visit on a scale from 1 to 10. To see if there is a significant difference between the grades of the Aryzon and the Galaxy S7, a Mann-Whitney U test is used. The descriptive values and the results of the Mann-Whitney U test are seen in Table 6. The mean value of the Aryzon is a lot higher than the handheld and according to the Mann-Whitney U test that difference is significant. This is shown quite well in the boxplot and histograms, seen in Figure 22.

Device	Min.	Max.	Mean	Std. Dev.	Mean Rank	Sum of Ranks	MWU	p-value
Aryzon	7	10	8,69	0,855	19,46	253,00	33,00	0,001
Handheld	6	10	7,47	0,990	10,20	153,00		

Table 6: The descriptives and Mann-Whitney U test of the Grade of the entire visit.

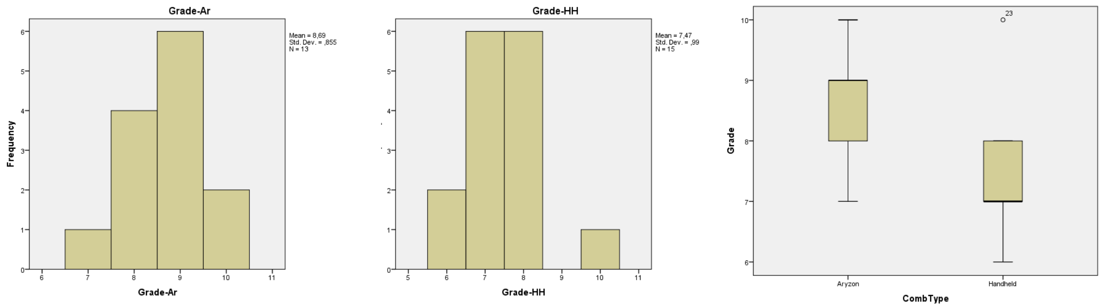


Figure 22: The left image shows the histogram of the Grade for the Aryzon, the lowest grade is a 7 and the highest a 10 and the highest frequency is 6. The center image shows the histogram of the Grade for the handheld, the lowest grade is a 6 and the highest a 10 and the highest frequency is 6. The right image shows the boxplot of the Grade for both the Aryzon (left) and the Galaxy S7 (right).

7.5 Discussion

The posed hypothesis states that the Aryzon is the preferred device, because of a higher immersion and if the usability and technical aspects are good enough. To measure this, data on the visitor experience is taken in the form of satisfaction, enjoyment and knowledge gained. Two statistical tests are performed on the retrieved data from the questionnaires. The t-test is used on the entirety of a visitor experience factor and the Mann-whitney U test is used on each individual question of a factor. Both tests indicate that the Aryzon is the preferred device on each factor. Looking at each factor, the users of the Aryzon are more satisfied about the content and how it is shown. From the enjoyment factor, they are more immersed, pleased and think the visit to the RSH is worthwhile. Aryzon users are also more motivated to learn about the RSH and think their visit was educational. Looking at the grade that users gave to their visit, this shows that overall the users of the Aryzon had a better visit than the users of the Galaxy S7.

However, the hypothesis also states that the Aryzon is preferred if the usability and technical aspects are good. The data obtained from the logging and observations indicate that users have a harder time using the Aryzon than the Galaxy S7. The observations don't show any indication of ergonomic issues. There are also notes that the AR objects the Aryzon shows are not always clear or on the correct location. All of these point towards a lack of usability and technological performance.

To conclude, even if the Aryzon does not perform well, it does perform good enough to create a better visitor experience compared to the Galaxy S7. Future research is needed to identify how much of an influence the performance has on the visitor experience.

8 Conclusion and Future Work

The goal of this study is to look at cheap AR solutions in the context of cultural heritage and see which is preferred by the visitors. These AR solutions are represented by the Aryzon, an AR cardboard device, and the Samsung Galaxy S7, a mobile handheld. To find the preference of the visitor, the two devices are compared on the visitor experience, which is divided in three measurable factors. These are the visitor enjoyment, visitor satisfaction of their visit and the knowledge gained by the visitor. With a between-subject design, the two AR devices are compared with visitors from the RSH and the results from the questionnaire show that the Aryzon is the preferred device. However, the observations made during the experiments and the data logged on the devices show that the Aryzon is harder to use and has some technical flaws. Still, the Aryzon is preferred by the visitors despite the drawbacks.

This research concluded that the Aryzon is the preferred device, however the results can be improved. To get these improvements, some changes can be made.

First of all, the two groups that participated are not identical. The group that used the Galaxy S7 had a lot of high school students that participated, because of a mandatory excursion. To get more accurate results, both groups should have a similar distribution in the age groups and motivations for their visit.

Second, the sample size was too small initially for the t-test on each individual question. Increasing the group sizes to 30 for both groups should have a positive influence on the results as it is a better representation of the population.

Another way to improve the test would be to put the applications more into the context of the tour. The current experiment felt more like an external addition to the house. The application should feel more natural and integrated with the tour to get more accurate results.

Finally, the results should not only focus on the entire visitor experience. The logging and observations already showed that the Aryzon is harder to use. A good addition to the results would be to test both AR devices on usability as well and find out how much this influences the visitor experience.

A different area of research that would be interesting, is to see how cardboard AR would perform in other areas where mobile handheld AR is used. In this study the Aryzon is preferred over the handheld, but would that always be the case? The Aryzon could also be compared with the Hololens. Both are AR HMDs, but a Hololens is expensive. Could the Aryzon give the same results as a Hololens, allowing those areas and applications to have a cheaper solution?

References

[1] Ericsson Consumerlab, 10 hot consumer trends, 2017

[2] Ding M, Augmented Reality in Museums, amt-lab.org, 2017

[3] Brancati et al, Usability Evaluation of a Wearable AR System for the Enjoyment of the Cultural Heritage, 11th International Conference on Signal-Image Technology and Internet-Based Systems (SITIS), Bangkok, Thailand, pp. 768–774, 2015

[4] Kerr et al, Wearable Mobile Augmented Reality Evaluating Outdoor User Experience, Proceedings of the 10th International Conference on Virtual Reality Continuum and Its Applications in Industry, Hong Kong, China, 2011

[5] Jeong J.H., Lee K.H., The physical environment in museums and its effects on visitors' satisfaction. Building and Environment, Volume 41, Issue 7, pp 963–969, 2006

[6] Pendit U.C., Zaibon S. B., Abubakar J. A., User Experience on Enjoyable Informal Learning via Mobile AR: Development and Evaluation, International Journal of Interactive Digital Media (IJIDM), Volume 2, Issue 2, pp 29-34, 2014

[7] Chang et al, Development and behavioral pattern analysis of a mobile guide system with augmented reality for painting appreciation instruction in an art museum, Computers & Education, Volume 71, pp 185-197, 2014

[8] Caggianese G., Neroni P., Gallo L., Natural Interaction and Wearable Augmented Reality for the Enjoyment of the Cultural Heritage in Outdoor Conditions, De Paolis L., Mongelli A. (eds) Augmented and Virtual Reality, Lecture Notes in Computer Science, Volume 8853, pp 267-282, 2014

[9] Leue M.C., Jung T., tom Dieck D., Google Glass Augmented Reality: Generic Learning Outcomes for Art Galleries, Tussyadiah I., Inversini A. (eds) Information and Communication Technologies in Tourism, 2015

[10] Lin A., Gregor S., Ewing M, Developing a Scale to Measure the Enjoyment of Web Experiences, Journal of Interactive Marketing, Volume 22, Issue 4, pp 40-57, 2008

[11] Dunleavy M., Dede C., Mitchell R., Affordances and Limitations of Immersive Participatory Augmented Reality Simulations for Teaching and Learning, Journal of Science Education and Technology, Volume 18, Issue 1, pp 7-22, 2009

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

[13] Billinghurst M., Clark A., and Lee G., A Survey of Augmented Reality, Foundations and Trends in Human-Computer Interaction archive, Volume 8, Issue 2-3, pp 73-272, 2015

[14] Reference on general AR in introduction Azuma R.T., A Survey of Augmented Reality, Presence: Teleoperators and Virtual Environments, Volume 6, Issue 4, pp 355-385, 1997

[15] Van Zijl I., Het ontstaan van het Rietveld Schröderhuis, Rietveld Schröderhuis, pp 15-31, 2017

[16] ICAR, https://www.icar.cnr.it/en/progetti/pon-orchestra/

[17] Aryzon, https://aryzon.com

[18] HoloKit, https://holokit.io

[19] Lenovo Mirage AR headset, https://www3.lenovo.com/us/en/jedichallenges

A Histograms and boxplots

These boxplots show if the distribution of the sample sets are similar. This is necessary for the Mann-Whitney U test.

Enjoyment

These show the plots of the questions related to the Enjoyment factor.

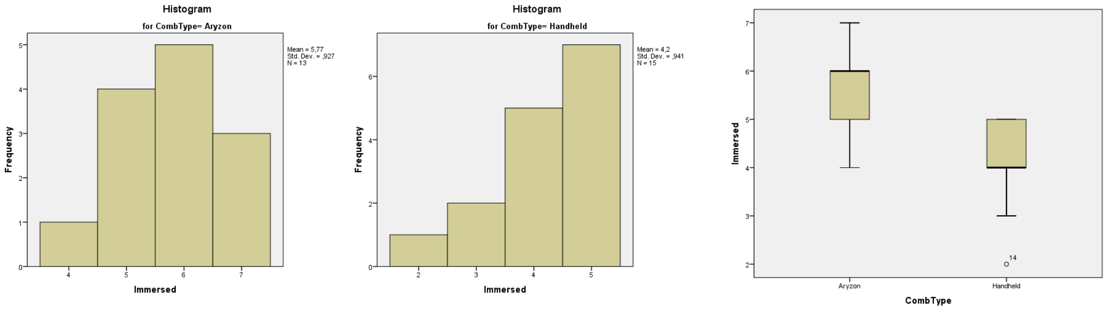


Figure 23: The left image shows the histogram of the Immersed value for the Aryzon. The center image shows the histogram of the Immersed value for the handheld. The right image shows the boxplot of the Immersed value for both the Aryzon (left) and the handheld (right).

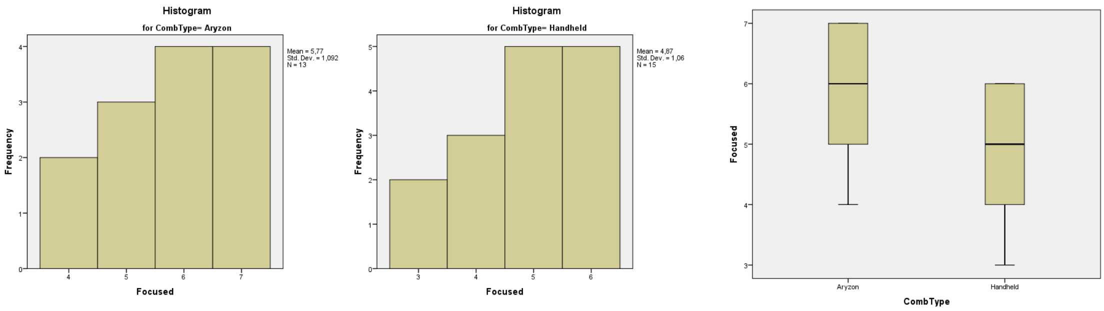


Figure 24: The left image shows the histogram of the focused value for the Aryzon. The center image shows the histogram of the focused value for the handheld. The right image shows the boxplot of the focused value for both the Aryzon (left) and the handheld (right).

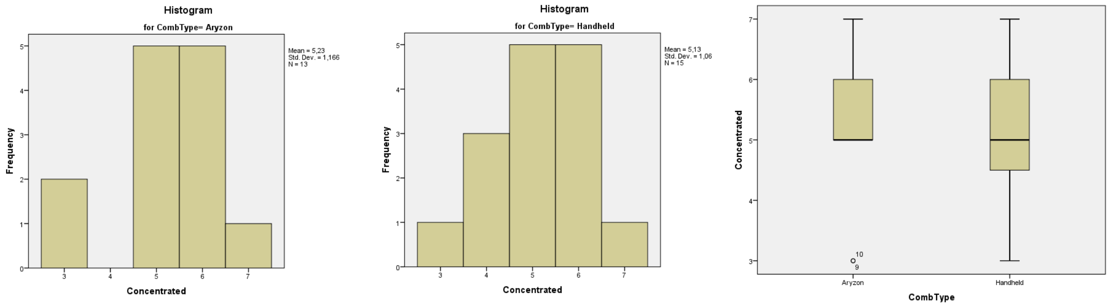


Figure 25: The left image shows the histogram of the concentrated value for the Aryzon. The center image shows the histogram of the concentrated value for the handheld. The right image shows the boxplot of the concentrated value for both the Aryzon (left) and the handheld (right).

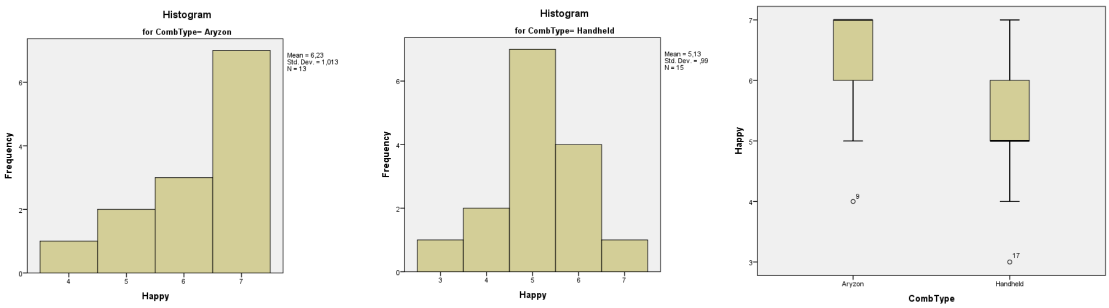


Figure 26: The left image shows the histogram of the happy value for the Aryzon. The center image shows the histogram of the happy value for the handheld. The right image shows the boxplot of the Immersed value for both the Aryzon (left) and the handheld (right).

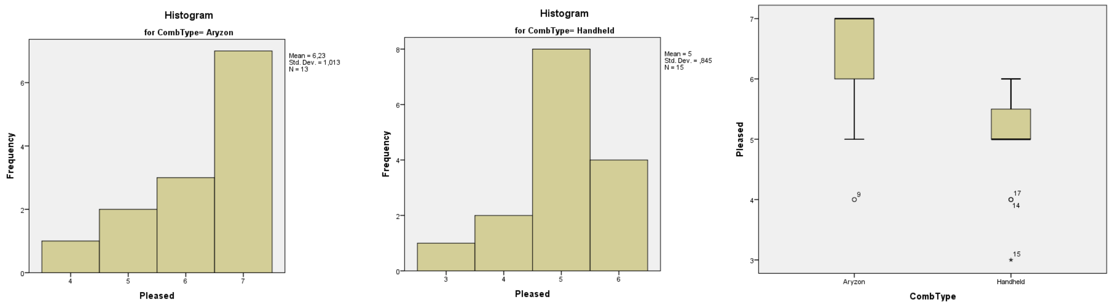


Figure 27: The left image shows the histogram of the pleased value for the Aryzon. The center image shows the histogram of the pleased value for the handheld. The right image shows the boxplot of the pleased value for both the Aryzon (left) and the handheld (right).

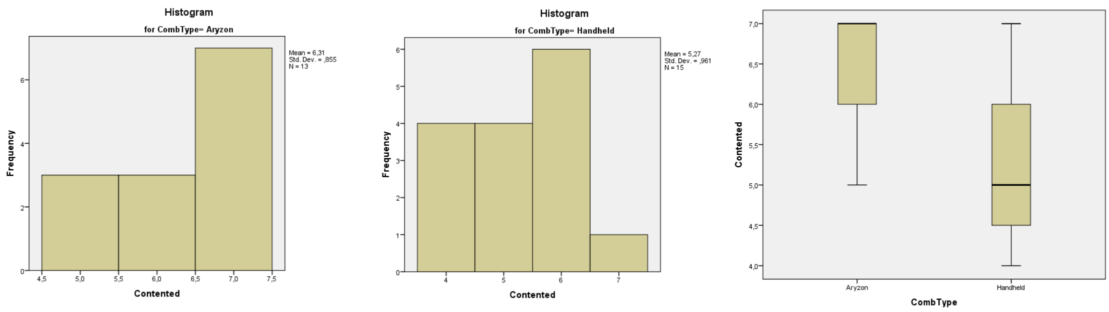


Figure 28: The left image shows the histogram of the contented value for the Aryzon. The center image shows the histogram of the contented value for the handheld. The right image shows the boxplot of the contented value for both the Aryzon (left) and the handheld (right).

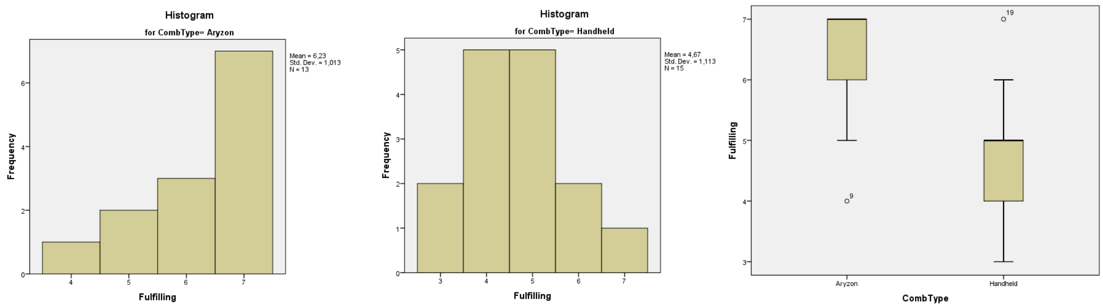


Figure 29: The left image shows the histogram of the fulfilling value for the Aryzon. The center image shows the histogram of the fulfilling value for the handheld. The right image shows the boxplot of the fulfilling value for both the Aryzon (left) and the handheld (right).

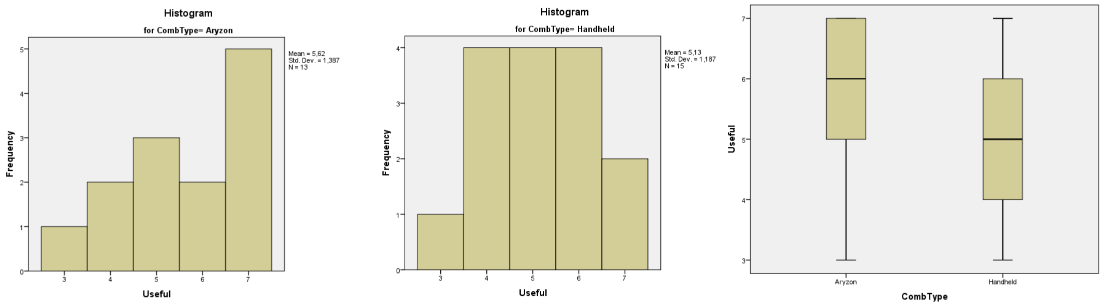


Figure 30: The left image shows the histogram of the useful value for the Aryzon. The center image shows the histogram of the useful value for the handheld. The right image shows the boxplot of the useful value for both the Aryzon (left) and the handheld (right).

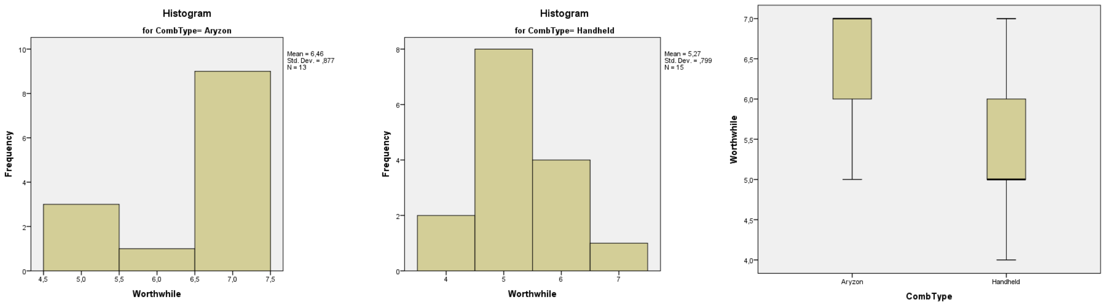


Figure 31: The left image shows the histogram of the worthwhile value for the Aryzon. The center image shows the histogram of the worthwhile value for the handheld. The right image shows the boxplot of the worthwhile value for both the Aryzon (left) and the handheld (right).

Satisfaction

These show the plots of the questions related to the Satisfaction factor.

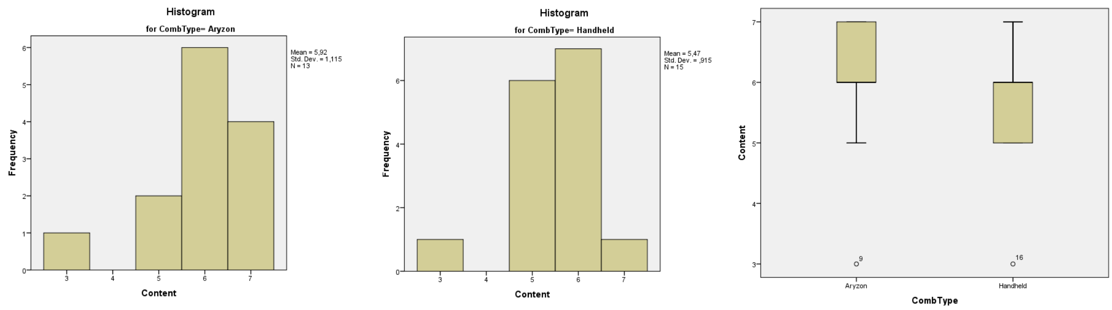


Figure 32: The left image shows the histogram of the content value for the Aryzon. The center image shows the histogram of the content value for the handheld. The right image shows the boxplot of the content value for both the Aryzon (left) and the handheld (right).

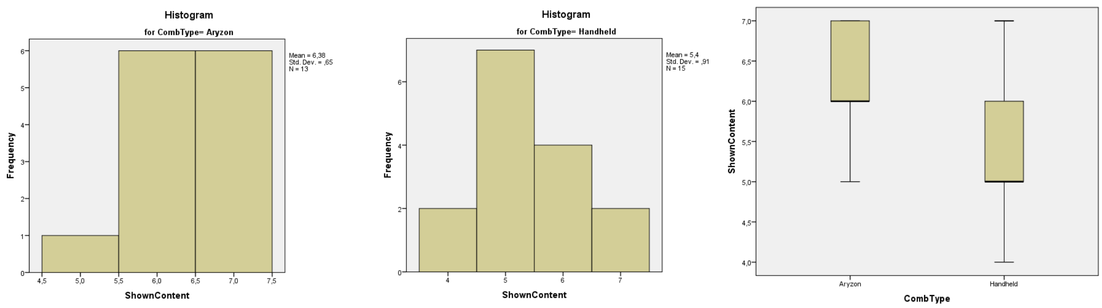


Figure 33: The left image shows the histogram of the showncontent value for the Aryzon. The center image shows the histogram of the showncontent value for the handheld. The right image shows the boxplot of the showncontent value for both the Aryzon (left) and the handheld (right).

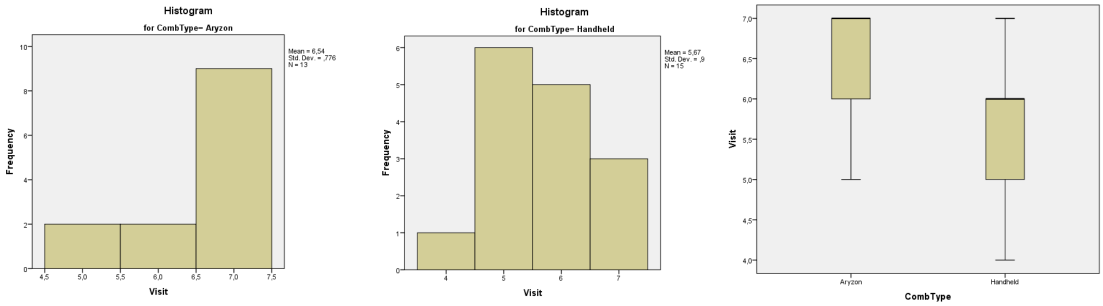


Figure 34: The left image shows the histogram of the visit value for the Aryzon. The center image shows the histogram of the visit value for the handheld. The right image shows the boxplot of the visit value for both the Aryzon (left) and the handheld (right).

Knowledge Gained

These show the plots of the questions related to the Knowledge factor.

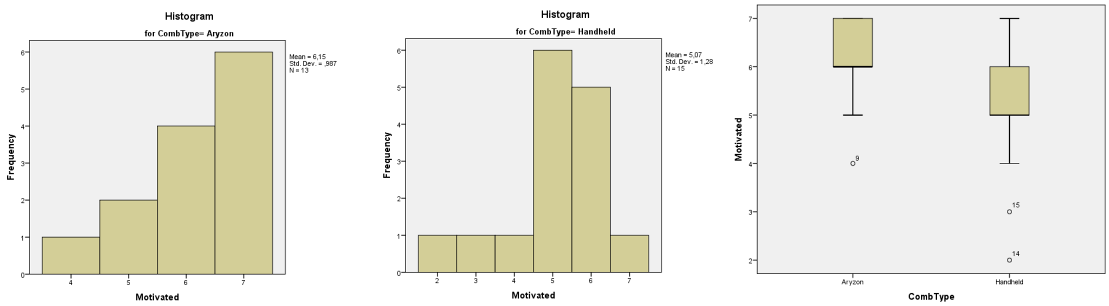


Figure 35: The left image shows the histogram of the motivated value for the Aryzon. The center image shows the histogram of the motivated value for the handheld. The right image shows the boxplot of the motivated value for both the Aryzon (left) and the handheld (right).

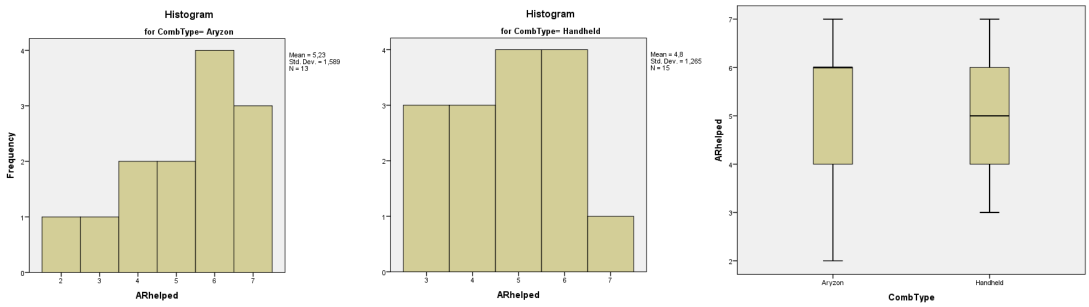


Figure 36: The left image shows the histogram of the arhelped value for the Aryzon. The center image shows the histogram of the arhelped value for the handheld. The right image shows the boxplot of the arhelped value for both the Aryzon (left) and the handheld (right).

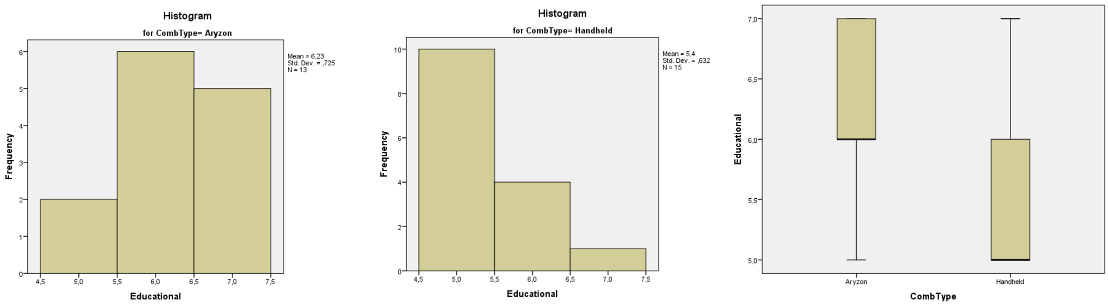


Figure 37: The left image shows the histogram of the educational value for the Aryzon. The center image shows the histogram of the educational value for the handheld. The right image shows the boxplot of the educational value for both the Aryzon (left) and the handheld (right).

B Questionnaire

This appendix contains the questionnaire that is used in the experiment. It is co-created with Joost Overbeek and thus contains questions for his study as well. The questions on demographic information about age, gender, visit, reason to visit and experience are used in this study. Furthermore the questions on Enjoyment, Satisfaction and Learning are used as well. Lastly, the final page is used in this research too. The questionnaire is found on the next page.

Visit Rietveld Schröder house

Information

What is your age

- ◯ 18-25
- ◯ 25-40
- ◯ 40-60
- ◯ 60+
- ◯ I prefer not to answer

What is your gender

- ◯ male
- ◯ female

What is your highest finished education

- ◯ High School/Middelbare school
- ◯ Community college/MBO
- ◯ Bachelor/HBO
- ◯ University Master
- ◯ Other
- ◯ I prefer not to answer

What is your current employment status?

- ◯ Student
- ◯ Part time (up to 34 hours/week)
- ◯ Full time (35+ hours/week)
- ◯ Retired
- ◯ Other

Have you visited the Rietveld Schröder House before?

- ◯ Never
- ◯ Once
- ◯ Twice
- ◯ Three or more times

What is your reason to visit the Rietveld Schröder House?

What is your experience with Virtual Reality (VR) and/or Augmented Reality (AR)?

- ◯ Never heard of it
- ◯ I know of it but have never tried it
- ◯ I have tried it
- ◯ I have used it more than a few times

Visit Rietveld Schröder house

Enjoyment

While visiting the Rietveld Schröder House I was immersed

Strongly disagree ○ ○ ○ ○ ○ ○ ○ Strongly agree

While visiting the Rietveld Schröder House my attention was focused

Strongly disagree ○ ○ ○ ○ ○ ○ ○ Strongly agree

While visiting the Rietveld Schröder House I concentrated fully

Strongly disagree ○ ○ ○ ○ ○ ○ ○ Strongly agree

While visiting the Rietveld Schröder House, I felt happy

Strongly disagree ○ ○ ○ ○ ○ ○ ○ Strongly agree

While visiting the Rietveld Schröder House, I felt pleased

Strongly disagree ○ ○ ○ ○ ○ ○ ○ Strongly agree

While visiting the Rietveld Schröder House, I felt contented

Strongly disagree ○ ○ ○ ○ ○ ○ ○ Strongly agree

Visiting the Rietveld Schröder House was fulfilling

Strongly disagree ○ ○ ○ ○ ○ ○ ○ Strongly agree

Visiting the Rietveld Schröder House was useful

Strongly disagree ○ ○ ○ ○ ○ ○ ○ Strongly agree

Visiting the Rietveld Schröder House was worthwhile

Strongly disagree ○ ○ ○ ○ ○ ○ ○ Strongly agree

Visit Rietveld Schröder house

Satisfaction

What did you think of the (exhibited) content of the Rietveld Schröder House.

Very unsatisfactory ○ ○ ○ ○ ○ ○ ○ Very satisfactory

What did you think of the way the (exhibited) content was shown.

Very unsatisfactory ○ ○ ○ ○ ○ ○ ○ Very satisfactory

What did you think of your visit to the Rietveld Schröder House.

Very unsatisfactory ○ ○ ○ ○ ○ ○ ○ Very satisfactory

Visit Rietveld Schröder house

Learning

I was motivated to learn about the Rietveld Schröder House.

strongly disagree ○ ○ ○ ○ ○ ○ ○ strongly agree

The augmented reality device helped me to learn more about the Rietveld Schröder House.

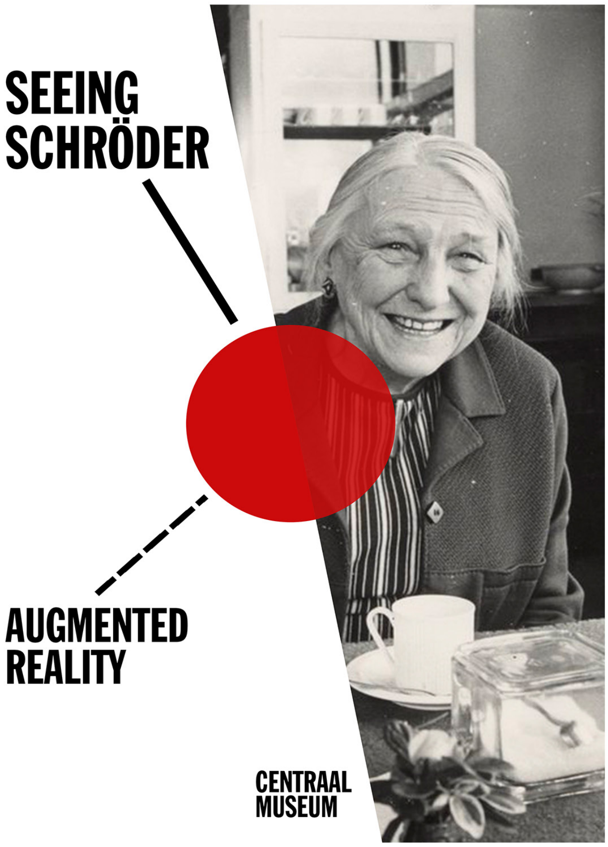
strongly disagree ○ ○ ○ ○ ○ ○ ○ strongly agree

The visit to the Rietveld Schröder House was educational/informative.

strongly disagree ○ ○ ○ ○ ○ ○ ○ strongly agree

Visit Rietveld Schröder house

Marker 1



This marker makes me curious

strongly disagree ○ ○ ○ ○ ○ ○ ○ strongly agree

This marker fits the situation

strongly disagree ○ ○ ○ ○ ○ ○ ○ strongly agree

This marker fits the house

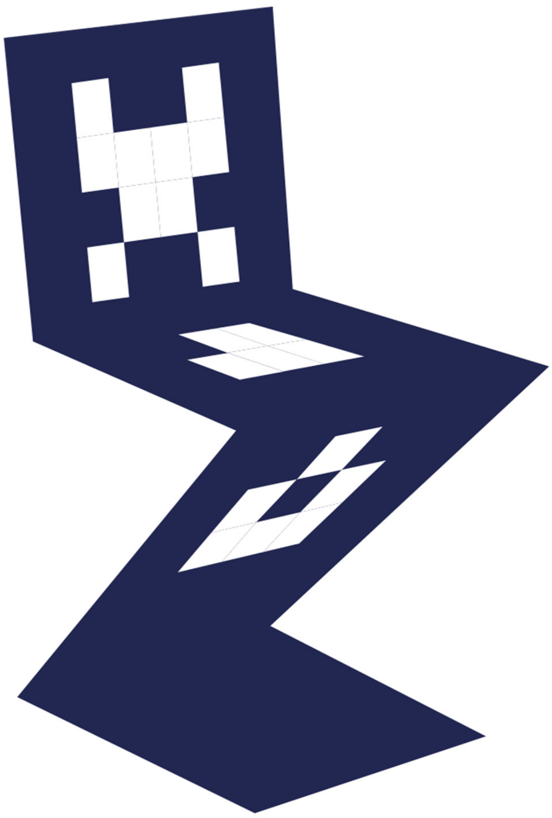
strongly disagree ○ ○ ○ ○ ○ ○ ○ strongly agree

I think the marker is attractive

strongly disagree ○ ○ ○ ○ ○ ○ ○ strongly agree

Visit Rietveld Schröder house

Marker 2



This marker makes me curious

strongly disagree ○ ○ ○ ○ ○ ○ ○ strongly agree

This marker fits the situation

strongly disagree ○ ○ ○ ○ ○ ○ ○ strongly agree

This marker fits the house

strongly disagree ○ ○ ○ ○ ○ ○ ○ strongly agree

I think the marker is attractive

strongly disagree ○ ○ ○ ○ ○ ○ ○ strongly agree

Visit Rietveld Schröder house

Augmented Reality

I felt comfortable in using the device

strongly disagree ○ ○ ○ ○ ○ ○ ○ strongly agree

Using the device was self-explanatory

strongly disagree ○ ○ ○ ○ ○ ○ ○ strongly agree

The device invites use

strongly disagree ○ ○ ○ ○ ○ ○ ○ strongly agree

The device makes me curious

strongly disagree ○ ○ ○ ○ ○ ○ ○ strongly agree

Using the device was fun

strongly disagree ○ ○ ○ ○ ○ ○ ○ strongly agree

Using the device was informative

strongly disagree ○ ○ ○ ○ ○ ○ ○ strongly agree

The experience is immersive

strongly disagree ○ ○ ○ ○ ○ ○ ○ strongly agree

Visit Rietveld Schröder house

Overall

What would you grade this visit?

1 ☆ ☆ ☆ ☆ ☆ ☆ ☆ ☆ ☆ ☆ 10

Would you recommend this visit to other people

- ◯ Yes
- ◯ Yes, but not to everyone
- ◯ No

Do you have any remarks or comments about the use of the augmented reality device?

Do you have any remarks or comments about the experiment?

Visit Rietveld Schröder house

Consent

- ◯ By checking this, I agree I was aware that I was allowed to stop during the experiment at any time for any reason
- ◯ By checking this, I agree the data from this survey is used anonymously for further development

I would like to get an update on the development of this project (please fill in email adress)