How Augmented Reality Affects the Learning Experience in a Museum

EMILIO LANDO
How Augmented Reality Affects the Learning Experience in a Museum

Emilio Lando
Lando@kth.se

Master’s Thesis at Mixed Reality Stockholm and CSC

Human Computer Interaction and
Master’s Program in Computer Science

KTH Supervisor: Mario Romero
KTH Examiner: Tino Weinkauf
Principal: Anders Hedman
Date: 2017-06-20
Abstract

It is well documented that Augmented Reality (AR) enhances and supports learning. Earlier research compares AR applications with existing methods. Published research typically focuses on one AR in general. Nevertheless, there are different ways of using AR. This paper gives further insight into how to use AR as a learning tool as part of a museum experience. It focuses on AR through smartphones, where the world is measured through the phone’s sensors and the virtual content is displayed on the device’s screen. This thesis presents the results of a comparative study between two types of AR: In-world space and On-screen space. In-world space AR renders the virtual content registered onto the physical exhibition. On-screen space AR renders the virtual content on the screen of the phone and uses the physical space of the exhibition as an index retrieval point. The discussion emerging from this study aims to aid the development and design of AR applications at museum settings, by giving curators better understanding of design options in AR spaces. Qualitative results suggest that In-world space benefits learning.
Sammanfattning

How Augmented Reality Affects the Learning Experience in a Museum

Emilio Lando
KTH Royal Institute of Technology
Stockholm, Sweden
lando@kth.se

ABSTRACT
It is well documented that Augmented Reality (AR) enhances and supports learning. Earlier research compares AR applications with existing methods. Published research typically focuses on one AR in general. Nevertheless, there are different ways of using AR. This paper gives further insight into how to use AR as a learning tool as part of a museum experience. It focuses on AR through smartphones, where the world is measured through the phone’s sensors and the virtual content is displayed on the device’s screen. This thesis presents the results of a comparative study between two types of AR: In-world space and On-screen space. In-world space AR renders the virtual content registered onto the physical exhibition. On-screen space AR renders the virtual content on the screen of the phone and uses the physical space of the exhibition as an index retrieval point. The discussion emerging from this study aims to aid the development and design of AR applications at museum settings, by giving curators better understanding of design options in AR spaces. Qualitative results suggest that In-world space benefits learning.

Keywords
Augmented Reality, Google Tango, In-world space, On-screen space, learning, short-term memory, interaction, museum, localization, smartphone.

1. INTRODUCTION
Augmented Reality (AR) is a technology to merge virtual content with the physical world. There is a Virtuality Continuum illustrating AR on a scale (see Figure 1). It ranges from the real environment to a virtual environment. AR is between these two in the continuum with virtual content integrated into the real environment [9]. The virtual content is registered on top of surfaces or in relation to them and should have a stable representation to be called AR. It should not jump around unintentionally or go through surfaces.

Figure 1: Virtuality Continuum [9].

A significant amount of published research addresses the AR application area of tourism and museums. Typically, the focus is on the technological aspect of AR or the end user [5]. With the focus on the end user, research generally compares one version of AR to other methods of interaction or receiving information. It typically tests if AR is beneficial for certain tasks. However, when AR is proven to benefit a certain area or task there is a general lack of research about best practices regarding how to use AR. There are different methods for using AR and this study presents a comparison between two such methods.

This study focuses on AR through smartphones. The world is measured through the smartphone’s sensors and virtual content is displayed on the device’s screen. In-world space and On-screen space are two methods of using AR for smartphones. The methods differ in how content is rendered. In-world space renders the virtual content onto the physical exhibition. On-screen space renders a virtual retrieval point onto the physical exhibition but renders the content on the screen of the phone.

The amount of information remembered by users can be a way of approximately measuring the learning effectiveness of methods displaying information. In-world space and On-screen space are used as such methods in this study. The integration of AR in learning is well documented. Santos et al. did a survey of 87 research papers in 2014 examining the learning effect of AR [10]. The report suggests that AR is beneficial for learning.

1.1 Motivation
The general lack of research comparing different methods of AR at museums suggests the need for this study. With the suggested benefit of AR in learning and the natural need for displaying information at a museum, learning is the focus for the comparison between In-world space and On-screen space. Short-term memory and recognition are the aspects in learning measured to examine the effectiveness of the two methods, thus giving further insight into design and use of AR at a museum.

1.2 Research Question
What are the effects of augmented reality on learning in the context of a museum when displaying information in “In-world space” versus in “On-screen space”, as approximated to measures of short-term memory recognition tasks?

2. BACKGROUND

2.1 Augmented Reality in Learning
The benefit of the integration of AR into learning can be explained with two theories about learning, Situated Learning and Constructivist Theory. Situated Learning suggests that
information conceived in a specific context enhances learning. The quality of learning is a result of interaction between people, places, objects and processes within a context [3]. The core of AR is content augmented onto an environment [9], thereby having context. User interaction is generally a part of an AR experience, further suggesting a benefit in learning with AR. Constructivist Theory explains that learning is based on what a person already knows and believes, which is shaped by the person’s prior experiences. It outlines five conditions which most likely enhances learning: 1) Embed learning within relevant environments; 2) Make social negotiation integral to the learning experience; 3) Provide multiple perspectives and multiple modes of representation; 4) Provide self-directed and active learning opportunities; and 5) Support and facilitate metacognitive strategies within the experience [6]. AR as a tool, fulfills several of those conditions. The enhancement of learning using AR is based on these two theories [6].

Research justifies the presumption that learning benefits from AR. It suggests that AR and digital displaying of information enhances learning. Santos et al. wrote a survey, in 2014, of research papers for AR in learning. It examined 87 articles since 2007 with the inclusion criteria: 1) The research paper must have at least a preliminary working ARLE (Augmented Reality Learning Experience) prototype; 2) The prototype should be applied to learning a new concept or skill; 3) The content should be relevant to kindergarten, primary and/or secondary education (However, content need not be tested on these target students); 4) The full research paper is publicly accessible; 5) The paper reports an effect size or provide means to calculate the effect size (reports both mean and standard deviation).

Effect size is a method of quantifying the size between two groups. Rather than explaining if something works or not, it explains how well it works in a range of contexts. Effect size can be directly converted into the overlap between two samples in terms of comparison of percentiles.

\[
\text{Effect Size} = \frac{\text{Mean of experimental group} - \text{Mean of control group}}{\text{Standard deviation}}
\]

An effect size is equivalent to a “Z-score” of a standard normal distribution. E.g. an effect size of 0.8 means that the score of an average person in the experimental group is 0.8 standard deviations above the average person in the control group, thereby exceeding 79% of the control group.

The survey indicated that ARLE has a significant effect on learning. The mean effect size was 0.56 for ARLE which is higher than the reported 0.35 for technology. A mean effect size of 0.56 for ARLE compared to non-digital learning methods means 71% of the group using an ARLE will perform better than the non-digital control group [10].

The novelty of AR might be a reason for the positive effect of AR for learning. This novelty could have a positive influence on the attitude and enjoyment of AR which increases the motivation for learning [13]. When the novelty fades, the positive effect might become less prominent.

A study conducted at Franklin Institute Science Museum focuses on learning at a museum. It compared AR as a mean of receiving information to existing practices currently used at the museum. The study showed that AR had a significant increase in learning in an informal setting such as a museum [14].

Research focuses on AR in regards to contextual content, interaction and visualization but there is a general lack of research about how to use AR. In-world space and On-screen space are two methods of displaying content in AR and this study compares them. Comparing these methods may give further insight into best practices for using AR as a tool for presenting information at museums.

2.2 Applications using AR in Museums

There are different types of content possible to augment in an AR application. CHESS framework and Skin and Bones are two AR-applications used at museums which augment a wide range of content [4][11]. The CHESS framework is a research project which aims to give museums the ability to create augmentation of their exhibitions. Skin and Bones is an application created for the Smithsonian National Museum of Natural History.

2.1.1 CHESS Framework

The CHESS framework aims to help museums make their collections more engaging, specifically for the younger and more digitally driven visitors. CHESS proposes to create a narrative-driven cultural adventure for the visitors with personalized storytelling of the exhibition [4]. CHESS follows a plot-based approach where the museum creates a story, based on different themes at the museum. The user follows scripted paths through the museum to trigger different events and scripts. The path tells a story by augmenting multimedia resources such as audio files, videos, pictures and so on, on different exhibitions at the museum. The user decides what content to receive and ignore by having a critical path and optional sub-paths, thus creating a personal experience [12]. With audio files, videos, pictures etc. it is possible to use already available content from the curators.

2.1.2 Skin and Bones

The Smithsonian National Museum of Natural History is a museum with an exhibition called “Bone Hall” featuring skeletons. Skin and Bones creates an augmentation of this exhibition by adding different media to the artifacts. The museum created media in form of videos, audio and virtual 3D models such as skin and flesh. The museum created the media. The skin layer augmented on the physical skeletons gives life to the exhibition [11].

The scope of this thesis requires the content augmented to be images and text. The content should complement the physical artifact. No 3D-models, or other content requiring a significant time to produce, will be used for this study.

2.3 AR techniques for museum environments

Museums typically display information in an indoor environment. This study will use AR technology for such environments. Indoor AR applications benefit from a more predictable movement of users than outdoor AR applications. Adding sensors, optical markers, location cameras and/or other localization technologies to important areas can prepare an environment for AR applications [1]. Indoor AR generally uses visual tracking with markers or visual features of the
environment [2]. A problem with visual features is the need for invariance. With changing light conditions, this is problematic [2]. Using visual markers is a stable visual tracking method for AR but it distracts from the user experience. To not overly distract users while displaying virtual content, it is desirable to hide or keep the markers as invisible as possible for the users [7].

New technologies use computer vision to remove the need for visual markers. Google Tango, for example, is a technology that has entirely removed the need for markers.

2.3.1 Google Tango

Google announced their Tango platform as a real-time location and mapping mobile AR platform in 2014. This technology solves the problems mentioned earlier with visual-based tracking. Google Tango uses computer vision to give the device knowledge about its location relative to the physical world around it. It uses motion tracking, area learning and depth perception to give an accurate localization relative to the physical world, removing the need for markers and visual features [8]. This study will use this technology for developing applications used to compare In-world space and On-screen space.

There are three core concepts making Google Tango possible: Motion Tracking, Area Learning and Depth Perception.

**Motion Tracking**

A Google Tango device uses motion tracking to keep track of its position in 3D-space. GPS-based localization solutions use GPS to locate the device in an outdoor environment. Google Tango instead keeps a relative location to its origin which corresponds to its location at the start of the service. Readouts from the device’s sensors update the pose (position and rotation) of the device continuously [8].

**Area Learning**

With Area Learning the device may recognize an area it has visited. The device can identify key features of the environment with data from a stereo camera, a depth sensor and a fish eye camera. By storing this data, the device can improve the trajectory in an environment, this is called relocalization and drift-correction (see Figure 2). Motion tracking alone becomes less accurate with time and drifts away from the device’s true trajectory. Area Learning corrects this inaccuracy, thus giving a more accurate position in 3D-space.

![Figure 2: Illustration of Google Tango drift-correction](image)

**Depth Perception**

With Depth Perception, a Google Tango device can understand distances. It uses one of three techniques: Structured Light, Time of Flight or Stereo Cameras. Structured Light and Time of Flight requires an IR-sensor while Stereo Cameras do not. The Tango API provides depth data as a point cloud where each point has an x-, y-, and z-coordinate and a confidence value representing the accuracy of the point [8].

3. METHOD

The purpose of this study was to give insight to museum curators about how to complement an exhibition with AR. In-world space and On-screen space are the two methods of displaying AR-content compared in this study. Short-term memory and recognition of AR-content were examined to measure learning for the two methods. Quantitative and qualitative data were collected to understand the differences between In-world space and On-screen space and give curators insight about design choices for AR at museums.

3.1 Study Design

3.1.1 Context

This study was conducted at the Museum of Far Eastern Antiquities in Stockholm. Two exhibition stations, Station A and Station B, had content connected to them. Station A was a statue of Ganesha, a Hindu god. Station B was a collection of Buddhist art and objects influenced by Greek culture.

The study did not replace the existing content at the exhibitions but complemented them with new information. This eliminated the risk of obtaining knowledge from other sources than the implemented AR prototypes, which was not the purpose of this study.

3.1.2 Prototype

With Google Tango is no need for markers or visual features which makes it a suitable AR-technology for AR-applications at a museum. A Google Tango smartphone called Lenovo Phab Pro 2 was used in this study which is the first commercially available smartphone with the Google Tango technology in Q4 of 2016. Two prototypes were implemented. Both prototypes displayed the same content.

**Prototype A**

Prototype A displayed the information in On-screen space. The user visited a station at the museum. The prototype augmented a virtual retrieval point onto the physical exhibition (see Figure 3). The device rendered the content on the screen of the device when the user touched the retrieval point. The content was then detached from the physical exhibition (see Figure 4).
Prototype B

Prototype B displayed the information in In-world space. It augmented content onto the physical exhibition. All content was related to the station where it was displayed. The user needed to aim the AR-device at that location to see the content (See Figure 5).

3.1.3 User Study

A user study was conducted to answer the research question with both quantitative and qualitative data. The study had within-group design and was counterbalanced. Every subject did one station with Prototype A and one station with Prototype B. The order of which prototype the subjects used first was evenly distributed between the subjects. Furthermore, the order of which stations the subject visited was also evenly distributed.

This generated four conditions. Three subjects executed each condition giving a total of twelve subjects. The subjects consisted of six women and six men. After visiting both stations the subject answered three surveys: one survey for each station and its content, and a user experience survey. Twelve subjects generated 24 results (twelve per prototype).

The focus group was adults (18+) without prior knowledge about the content displayed at the exhibition. Prior experience with AR was deemed unimportant due to the low interaction and complexity of the application. However, a requirement was that the subjects had experience with smartphones. They were required to be familiar with camera feeds and regular handling of such devices. Quantitative and qualitative data were collected via the use of surveys and observations.

Within-group design ensured a better qualitative comparison between the two prototypes. It assured a direct comparison between the prototypes for every subject. The result for both prototypes reflected each subject’s attitude and experience with the technology, thus making for a better comparison.

The subject received minimal information about the survey and content prior to the study. This reduced the risk of adapting to the study and actively making trade-offs to perform better on the surveys. The subjects were told that there will be surveys about what they see and experience, thus ensuring the need to visit both stations before answering any survey. Time between experiencing a station and answering the survey differed, which may affect the results. The subjects answered the survey connected to the first station they visited first to minimize this effect.
How Augmented Reality Affects the Learning Experience in a Museum

The content was consistent for each station in both prototypes. Between station A and station B, the content was inconsistent and there was no direct comparison between the survey results of station A and station B. The focus was on the overall comparison between Prototype A and Prototype B for both stations. The results also consists of a comparison between the two prototypes split up between the two stations.

3.1.3 Augmented Content

The augmented content were images with and without text. Both stations displayed related content for that exhibition. Each survey contained ten questions with a focus on both features in the images and the text in the images.

3.2 Pilot Study

Eight subjects took part in a pilot study. The study design was similar to the main study. Changes between the pilot study and the main study improved the quality of the quantitative data. Increasing the number of questions per survey generated more data for the quantitative analysis. An "I don’t know" option was added. This reduced the effect of luck due to the subjects no longer being required to guess. The survey results per station from the pilot study were discarded. However, the qualitative results were analyzed and used. The prototypes did not change between the pilot study and the main study.

In the pilot study, each subject visited one station and answered the survey related to it before visiting station two. It was observed that some subjects adapted and approached the second application and station differently. Thus, the change to make the participants visit both exhibitions before answering any survey in the main study was necessary to avoid subjects who adapted to the task.

4. RESULTS

The result consists of the analysis of quantitative and qualitative data. The focus is on the comparison between On-screen space and In-world space. Quantitative data is used to measure short-term memory and recognition. Qualitative data is used to further understand the quantitative data and give insight for museums about design choices regarding AR applications at museums.

4.1 Quantitative results

The analysis of the quantitative data indicates that there is no statistically significant difference between the two methods. The overall result is spread out for both On-screen space and In-world space, ranging from zero correct answers to nine correct answers (see Figure 8). It is clear that the results have a high deviation and are evenly distributed between the best and the worst result. In-world space is concentrated around the lower results (2-4) while On-screen space is concentrated at the mid ranges (4-6).

A comparison of the result between the two methods indicates that there is no significant difference between them. Figure 9 shows that the total amount of correct answers for In-world space was 48 while it was 49 for On-screen space. This
gives means of 4/10 and 4,1/10 for In-world space and On-screen space respectively.

**Figure 9:** The total amount of correct answers from In-world space and On-screen space.

The analysis for each station shows similar results. For station A, the two methods are close in number of correct answers and show similar distribution and deviation (see Figure 10). However, there is a slight difference in the amount of correct answers with In-world space having 31 and On-screen space having 26. Figure 11 illustrates the same condition but for station B. The distribution and deviation are similar for this condition as well. For station B, there is an inverse difference between the two methods. The amount of correct answers are 17 for In-world space and 23 for On-screen space.

As seen in Figure 9, the total amount of correct answers for In-world space differs from On-screen space only by one. This corresponds to around a two percent difference in favor of On-screen space. However, for Station A, the difference is five in favor of In-world space, which is a 19 percent difference. For Station B, the difference was six in favor or On-screen space which is 35 percent. The small sample size for each station (six subjects per condition) may be a reason for this difference and conclusions are difficult to draw.

**Figure 10:** In-world space vs. On-screen space on Station A.

**Figure 11:** In-world space vs. On-screen space on Station B.

Time spent on each station may affect the outcome of short-term memory and recognition. Table 1 shows the time spent on each application per subject. The average time spent on the In-world space application was 2 minutes and 23 seconds while it was 1 minute and 52 seconds for the On-screen space application. The difference in correct answers between In-world space and On-screen space was almost nonexistent. Although, the time spent using the In-world space application differs from the time spent on the On-screen space application by half a minute.

**Table 1.** Time spent on the two applications for each subject.

<table>
<thead>
<tr>
<th>In-world</th>
<th>On-screen</th>
</tr>
</thead>
<tbody>
<tr>
<td>02:05</td>
<td>01:32</td>
</tr>
<tr>
<td>01:30</td>
<td>00:40</td>
</tr>
<tr>
<td>00:51</td>
<td>00:31</td>
</tr>
<tr>
<td>02:35</td>
<td>01:52</td>
</tr>
<tr>
<td>01:35</td>
<td>01:15</td>
</tr>
<tr>
<td>01:57</td>
<td>02:02</td>
</tr>
<tr>
<td>02:15</td>
<td>01:35</td>
</tr>
<tr>
<td>03:54</td>
<td>02:45</td>
</tr>
<tr>
<td>04:45</td>
<td>02:43</td>
</tr>
<tr>
<td>01:27</td>
<td>02:20</td>
</tr>
<tr>
<td>01:57</td>
<td>01:39</td>
</tr>
<tr>
<td>03:47</td>
<td>03:33</td>
</tr>
</tbody>
</table>

After the subjects finished the surveys for both stations, they answered a survey about the experience, comparing In-world space and On-screen space. Figure 12 and Figure 13 show that the In-world space application was more enjoyable than the On-screen space application. The mean value for On-screen space is 3.2 with a standard deviation of 0.93 while the mean value for In-world space is 3.8 with a standard deviation of 0.75. Research shows that enjoyability and engagement are beneficial for learning. Therefore, this difference can affect the learning outcome from the application.
How Augmented Reality Affects the Learning Experience in a Museum

The Royal Institute of Technology, June 2017, Stockholm Sweden

Figure 12: Rating of On-screen space application.

Figure 13: Rating of In-world space application.

The ease of use of the application could also affect the result. A goal when implementing the prototypes was to keep the applications as easy to use as possible, thus, limiting the conditions that would affect the outcome of the study to the AR-methods as much as possible. Comparing Figure 14 and Figure 15 suggest that the goal was fulfilled. Both applications were considered to be easy to use and there was no significant difference between the two applications with regards to ease of use.

Figure 14: Rating of the ease of use of the On-screen space.

Figure 15: Rating of the ease of use of the In-world space.

Each subject answered a question about which application they preferred and why. In-world space was preferred by 13 subjects while six subjects preferred On-screen space. One subject had no preference (see Figure 16).

Figure 16: The number of subjects preferring each application

4.2 Qualitative results

The qualitative data is from 20 subjects. The differences between the pilot study and the main study only affected parts of the quantitative data. Therefore, all collected qualitative data is used.

To further understand differences between In-world space and On-screen space, the subjects answered questions about the two applications, thus giving subjective and qualitative substance to the result.

Each subject answered which application they preferred and why. Enjoyment and engagement benefit learning and were recurring reasons for the preferred application. The subjects said more often that In-world space had this effect than On-screen space. This suggests that In-world space may be better for this task than On-screen space.

"I preferred In-world space since I could explore for myself, it was also more fun trying to find all the available info. The on-screen space blocked the entire camera and felt more like an E-book." - Subject 3
"The on-screen experience took away from the museum I think, you looked at the screen reading text, instead of looking at the actual objects. The In-world space application enhanced the experience, so I looked at the museum object and it was complemented with extra images." - Subject 4

"Boring to stand on one place with On-screen" - Subject 16

Comfortability and physical exhaustion were aspects that several subjects noted regarding the applications. This was not anticipated before the study. Subject 10 pointed out that the use of the In-world space application could be a problem for people with bad knees. Keeping this in mind is important when developing and using AR at a museum. Some quotes from the study follow.

"On-screen was better since it was easier and more comfortable to use." - Subject 18

"I prefer On-screen because it doesn't require me to stand (or squat.) at a specific location. This is a big advantage when it's crowded and for comfort." - Subject 1

"On-screen makes so that you can step aside and keep reading on your screen. I had the sun in my eyes at the exhibition but I could avoid this by moving to the side with On-screen." - Subject 7

The In-world space application had content located around the exhibition both above and below eye level of a regular sized person (1.5 to 1.8 meters). This, required subjects to stretch or bend to be able to get close to the content. This physical requirement was troublesome and several subjects noted this. Subjects 1, 4, 6 and 10 are some of the subjects that had concerns about the required bending or squatting. Subject 7 and 18 did not specifically talk about the trouble of bending but noted that the On-screen space application was more comfortable and the experience was better overall.

When asked what the subjects liked most and least about the two applications, In-world space was considered to be engaging and gave the user appreciated exploring possibilities. The physical requirements were the most noted negative aspect of the application.

"I liked that I need to go and look at the actual objects", In-world - Subject 9

"Cool way of learning about history. I liked that you had to walk around the statue to find more pictures with content." - In-world - Subject 10

"Having to bend over far to the ground to see", In-world - Subject 13

The responses for what was most and least liked about On-screen space also reinforced earlier stated results. The enjoyment of On-screen space was due to the ease of use, effectiveness and how comfortable it was. It was least liked because it was boring compared to In-world and it reduced the AR-experience.

"Clearly see the text and images. Hard to miss an image", On-screen - Subject 9

"Not having to hold the phone still while reading. It also allows you to read in a comfortable position rather than having to aim the camera at the object you are reading about", On-screen - Subject 14

"Can't get the AR experience since there is a screen blocking the view", On-screen - Subject 11

"Boring to stand on one place", On-screen - Subject 16

Some subjects suggested that a combination of the two applications would be preferable. The suggested application would have all content located as the In-world space application, but the user could view the content in On-screen space if the user desired it. This would give the user a more personalized experience and would help when using the application at a crowded place. It would also keep the appreciated engagement and exploring part of the In-world space, which aids learning, but reduce the physical requirements.

"Maybe there could be a mix between on-screen space and in-world space in which the images and texts are placed by the objects so that you can read them augmented reality style, but if you are far away and don't want to walk up to it you can click the image in world space and bring it up onto screen space..." - Subject 6

"I think that it should be possible to click on an image in the in-world space to get it on the screen. That way I don't have to sit down in an angle to read it. I think that a combination of them both would be best." - Subject 9

"A combination of the two applications would be really enjoyable. Having an in-world space with a possibility of clicking on a slide in the on-screen space would have been nice." - Subject 11
5. DISCUSSION

This study compared In-world space and On-screen space as AR-methods used to display information at a museum. The focus was on short-term memory with quantitative data. To further explain the quantitative data and give more insight into the different AR-methods, this study also gathered qualitative data.

The quantitative data overall had a high standard deviation for both In-world space and On-screen space. This gives little indication of the expected performance and relation between the two methods. The deviations also indicate that there is a significant uncertainty to the study. Conclusions about the relation between the method and the number of correct answers are difficult to draw.

The design of the study may be the reason for this outcome. Twelve subjects seem to be too few for this study. More subjects would most likely reduce the standard deviations and give a better indication of the performances of In-world space and On-screen space. It would reduce the impact of edge values such as zero and nine correct answers which were both present in this study. Removing edge cases is a method for reducing a standard deviation. However, due to the high spread in this data set, it is difficult to understand which are edge cases and which are not.

The mean value of number of correct answers for both methods were close to each other (4 for In-world and 4.1 for On-screen). However, the high standard deviations suggest that the similarity of the means may be a coincidence.

When comparing In-world space and On-screen space one station at a time, the results were inconsistent. In-world space favored Station A while On-screen space favored Station B. The sample sizes for this data are small, six subjects per condition. Edge cases with such small sample sizes have a significant impact on the result. A subject who performs well on both stations affects the overall result accordingly. E.g. subject 12 performed well on both stations (9 on In-world Station A and 7 on On-screen Station B). This subject will not favor any method significantly in the overall comparison. With focus one station at a time, this subject will only be reflected in the result for one method. This generates a significant uncertainty to the result. This aspect combined with a small sample size, makes this station specific comparison irrelevant.

More content and more questions on the surveys is another solution to avoid low significance if the study is to be repeated. If the questions on the surveys were in greater numbers, the effects of the two methods on short-term memory could be more noticeable. This could give a better indication on how the two methods fare against each other.

The time spent on each specific prototype may affect learning. The average time spent on the In-world space application was 2 minutes and 23 seconds while it was 1 minute and 52 seconds for the On-screen space application. With the similar results between the two methods this may indicate that the On-screen space is more effective and/or easier to use than In-world space. However, the need for walking between the content and exploring the station can explain the longer time spent on the In-world space application. The On-screen space application only required swiping between images.

The qualitative data gave a clear indication on how the two methods compared to each other. In-world space was more enjoyable and more engaging than the On-screen space application. This is also strongly indicated from the combination of the quantitative results presented in Figure 12, 13 and 16. These aspects influence learning in a positive way. With a bigger sample size and more questions, these aspects could be a difference factor in a quantitative comparison with regards to short-term memory. It could give In-world space a better learning outcome with regards to short-term memory.

The main advantage On-screen space had over In-world space was comfortability and less physical requirement such as bending and stretching. This will most likely affect learning and the experience of an application when used for a longer period of time. Physical exhaustion and uncomfortable use distract the users from the task which reduces learning. It would also affect the availability of the application due to potential physical restrictions of visitors at a museum. However, this aspect is somewhat related to the design of the In-world space application and not the methods themselves. It is possible to achieve In-world space without requiring the subject to squat to the floor and stand in uncomfortable positions. There will still be more physical requirement when using In-world space but positioning the content more strategically may reduce the physical exhaustion. This would reduce the impact of the main aspect that gave On-screen space a better experience.

On-screen space was noted to be more comfortable not only due to the absence of squatting and stretching. For example, being able to get the content in a crowded area and read the content in a location and position of your own choosing is sometimes preferable. This may be the reason why the combination of the two methods got mentioned by some subjects. This combination may incorporate the best aspects of both In-world space and On-screen space. The engagement and enjoyment of walking around and exploring the exhibition will be present, and the possibility to read the content comfortably and at the user’s own conditions. Considering this design choice for an AR-application for museums may be beneficial.

6. CONCLUSIONS

This study compared AR in In-world space and On-screen space. The quantitative results gathered showed few indications on how the two methods compared against each other and consisted of data with high deviation.

The qualitative data suggests that In-world space have more aspects that would benefit learning. With a more extensive quantitative study, reducing the standard deviations and the effect of edge cases, these aspects might become noticeable. To be able to draw conclusions about the difference between In-world space and On-screen space with regards to short term memory a more extensive quantitative study is required. Future research is needed.

7. FUTURE RESEARCH

I suggest a more extensive quantitative study to further understand the differences between In-world space and On-screen space with regards to learning. To reduce the effect of edge cases and reduce the deviation, more subjects and content with more questions is suggested. This would give a better
indicating how the two methods perform in relation to each other.

Changing the design of the prototypes used in this study is also suggested. The effect of physical exhaustion and uncomfortability related to the In-world space had an impact on the experience which can affect learning and short-term memory. The placement of the content needs more thought to reduce the aspects affecting the result that are not strictly related to the two methods.

8. ACKNOWLEDGEMENTS
This study is supported and sponsored by the project Mixed Reality Stockholm and specifically the Museum of Far Eastern Antiquities.

REFERENCES


