

Article

Users' Perceptions Using Low-End and High-End Mobile-Rendered HMDs: A Comparative Study

M.-Carmen Juan ^{1,*} , Inmaculada García-García ², Ramón Mollá ¹  and Richard López ¹

¹ Instituto Universitario de Automática e Informática Industrial, Universitat Politècnica de València, C/Camino de Vera, s/n, 46022 Valencia, Spain; rmolla@dsic.upv.es (R.M.); ricloful@inf.upv.es (R.L.)

² Departamento de Sistemas Informáticos y Computación, Universitat Politècnica de València, C/Camino de Vera, s/n, 46022 Valencia, Spain; ingarcia@dsic.upv.es

* Correspondence: mcarmen@dsic.upv.es; Tel.: +34-963-879-720

Received: 10 January 2018; Accepted: 9 February 2018; Published: 13 February 2018

Abstract: Currently, it is possible to combine Mobile-Rendered Head-Mounted Displays (MR HMDs) with smartphones to have Augmented Reality platforms. The differences between these types of platforms can affect the user's experiences and satisfaction. This paper presents a study that analyses the user's perception when using the same Augmented Reality app with two MR HMD (low-end and high-end). Our study evaluates the user's experience taking into account several factors (control, sensory, distraction, ergonomics and realism). An Augmented Reality app was developed to carry out the comparison for two MR HMDs. The application had exactly the same visual appearance and functionality for both devices. Forty adults participated in our study. From the results, there were no statistically significant differences for the users' experience for the different factors when using the two MR HMDs, except for the ergonomic factors in favour of the high-end MR HMD. Even though the scores for the high-end MR HMD were higher in nearly all of the questions, both MR HMDs provided a very satisfying viewing experience with very high scores. The results were independent of gender and age. The participants rated the high-end MR HMD as the best one. Nevertheless, when they were asked which MR HMD they would buy, the participants chose the low-end MR HMD taking into account its price.

Keywords: mobile-rendered HMDs; head-mounted viewers; headsets; head-mounted displays; Augmented Reality; low-end mobile-rendered HMDs; high-end mobile-rendered HMDs; Samsung Gear VR; Google Cardboard; user's perceptions; comparative study

1. Introduction

User satisfaction with an Augmented Reality (AR)/Virtual Reality (VR) 3D app depends on the quality of the app, both for its content as well as its quality and realism. However, the graphical quality does not exclusively depend on the app graphics. The device that is used to output the 3D images also plays a key role [1].

Mobile devices have proliferated significantly in recent years and they are now available to a large majority of citizens. Current mobile devices have power that was unimaginable just 5 years ago. There are two reasons for the success of mobile devices: the hardware that is available and the tools that are available for programming. Mobile devices not only include fast CPUs, large displays, cameras with 16 MP, graphics acceleration, compass/accelerometers, GPS sensors or gyroscopes but they can also include depth and motion tracking sensors.

With regard to AR, it is possible to have AR using different types of devices, e.g., video see-through or optical see-through HMDs (Head-Mounted Displays). For video see-through HMDs, it is possible to have AR using VR HMDs or HMDs that are specific to AR. The main difference between the two

types is that VR HMDs do not include cameras for capturing the real world. With HMDs, users can perform tasks as “naturally” as they would do in everyday life [2]. The possibility of combining HMDs or Head-Mounted Viewers with smartphones and have an AR platform has opened a new niche for research and the market. According to Boos et al. [3], current VR HMDs fall into two device classes: (i) Tethered HMDs, HMDs that are tethered to powerful desktops, such as the Oculus Rift, HTC Vive and Sony PlayStation VR; (ii) Mobile-rendered HMDs (MR HMDs), self-contained, untethered HMDs that run on mobile phones slotted into head mounts, e.g., Google Cardboard, Samsung Gear VR, Google Daydream View, or Zeiss VR One. These mobile-rendered HMDs are screenless viewers made of cardboard or plastic and their cost is under 100 USD. According to several reports, the AR software market was valued at 2.13 billion USD in 2016 and is expected to reach 35.22 billion USD in 2022, with an annual growth of 57.36% [4]. The mobile AR market is expected to reach 79.77 billion USD by 2022, at an annual growth rate of 69.85% between 2016 and 2022 [5]. AR/VR headsets are expected to grow from 10 million units in 2016 to 100 million units in 2021, with an annual growth of 172.9% for AR headsets and an annual growth of 48.7% for VR headsets (the International Data Corporation (IDC) report). The major reasons for this growth are the increasing use of smart portable devices and the ease of implementation of AR applications. Versatility, reliability, ease of integration and price have also been identified as drivers of this growth.

It is worth mentioning that the following terms are currently being used to refer to the new devices of this type: HMDs, head-mounted viewers, headsets, or mobile-rendered HMDs. For consistency, the term “mobile-rendered HMDs/MR HMDs” will be used in the rest of the paper to refer to devices of this type.

Comparative studies about VR/AR users’ experiences among the different available MR HMDs will help researchers and users to choose the most appropriate device depending on its required use. In this line, the main objective of this paper is to compare a low-end MR HMD with a high-end MR HMD. Our first hypothesis was that the user’s experience using the two MR HMDs would not offer statistically significant differences. Our second hypothesis was that there would be statistically significant differences for the ergonomic factor in favour of the high-end MR HMD. To achieve the main objective and to check the two hypotheses, we developed an AR app and adapted it to each MR HMD. The app has exactly the same appearance and functionality for both MR HMDs. A study involving 40 adults compared the two MR HMDs and evaluated the user’s experience.

The paper is organized as follows. Section 2 briefly reviews related previous work. Section 3 describes the technical characteristics of the app developed. Section 4 presents details of the study. Section 5 reports on the results. Section 6 discusses the results and Section 7 presents our conclusions and future lines of research.

2. Background

2.1. Mobile-Rendered Head-Mounted Displays

This subsection mentions several of the current low-cost MR HMDs. Google Daydream View (<https://vr.google.com/daydream>) is a soft cloth VR screenless viewer that is compatible with many phones. It comes with its own Wii-like motion controller for manipulating objects and menus in VR. There is no need for adjusting lenses, no fiddling with pairing procedures and no ports or cables to plug in. The phone automatically pairs with the headset and adjusts the screen to the correct position. Tracking relies entirely on the phone. It is made of plastic and has a soft finish, which gives the headset a warm and ergonomic appearance.

Merge VR Goggles (<https://mergevr.com/goggles>) allow any recent models of Android and iOS smartphones to be used. This MR HMD does not add any additional sensors or electronics. However, the sides can be opened to access the smartphone ports.

Zeiss VR One (<http://vrone.us>) includes high quality lenses without lens distortion. It does not have adjustable interpupillary distance.

Google Cardboard (<https://vr.google.com/cardboard>) has an open design kit that anyone can make at home just by following a few instructions. It is one of the cheapest MR HMDs. It is representative of all the cheap clonic smartphone-based devices for AR/VR.

Samsung Gear VR (<http://www.samsung.com/global/galaxy/gear-vr>) was developed by Samsung Electronics in collaboration with Oculus. Any compatible Samsung Galaxy device works as the headset display and processor. The Samsung Gear VR unit acts as the controller. It contains a custom inertial measurement unit for rotational tracking. It connects to the smartphone via micro-USB. The Samsung Gear VR also includes a touchpad and back button on the side as well as a proximity sensor to detect when the device is on.

The rest of the low-cost MR HMDs are made of plastic with soft parts for ergonomic adaptation to the user's face. All of them are screen-less devices that require a mobile phone to support the AR/VR experience.

According to the technical features of the analysed devices, there are two differentiated MR HMDs: low-end MR HMDs represented by Google Cardboard and high-end MR HMDs represented by Samsung Gear VR. In this study, we use a Google Cardboard V2 (see Figure 1a) and a Samsung Gear VR Innovator Edition (SM-R320), see Figure 1b. Table 1 shows the characteristics of these two devices.



Figure 1. The MR HMDs used in our study: (a) Google Cardboard V2; (b) Samsung Gear VR Innovator Edition (SM-R320).

Table 1. Characteristics of Google Cardboard and Samsung Gear VR.

	Google Cardboard V2	Samsung Gear VR Innovation Edition (SM-R320)
Field of view	80°	96°
Pixel density	10	12
Building material	cardboard	plastic
Phone lock mechanism	Velcro	all-round clips
Supported mobile devices	A wide variety (screens from 4.0"–6.0")	Galaxy Note 4 For other models: Galaxy Note 5, S6, S6 Edge, S6 Edge+, S7, S7 Edge, Note 8, S8, S8+
Weight	96 g	379 g
Wearing comfort	none	foam
Control and sensors	button	volume key, touchpad, back key, focus adjustment wheel and proximity sensor
Adjustment to the user's head	No. The user's own hands or an external head wand is required	Yes

2.2. Previous Works with Mobile-Rendered HMDs

MR HMDs under 100 USD (e.g., Google Cardboard) are bringing VR and AR to the masses as the press and several authors have pointed out (e.g., [6]). These MR HMDs have already been used for a variety of purposes. For example, Google Cardboard has provided different VR experiences such as geographic tours in Machu Picchu and the Great Barrier Reef; another example of VR has been experienced by MacDonald's customers in their Happy Meal boxes [6].

Samsung Gear VR has also been used for different purposes. For example, Mosadeghi et al. [7] evaluated the acceptability and feasibility of VR in a diverse cohort of hospitalized patients using this MR HMD. Their study involved 30 participants, who used four different VR apps. A total of 75% of the sample believed that the experience could reduce pain by means of distraction. The patients were split on the comfort level; 52% found the device comfortable and 48% found it uncomfortable. Younger patients were more willing to participate. In a second study of the same group [8], 100 patients were involved (50 subjects in the VR cohort and 50 subjects in the control group). The pain reduction in the VR cohort was significantly greater than in the control group (means = -1.3 vs. -0.6 points, respectively; $p = 0.008$). No adverse events were reported from VR.

Steed et al. [9] used Samsung Gear VR and Google Cardboard in a study about presence and embodiment within VR in uncontrolled settings. The users of these two MR HMDs were invited to download and run an app that presented a scenario where the participant would sit in a bar watching a singer. Each participant saw one of eight variations of their scenario. From their results, the authors concluded that a self-avatar had a positive effect on the self-report of presence and embodiment and that the singer inviting the participant to tap along had a negative effect on the self-report of embodiment. That study, did not analyse the differences between the two MR HMDs.

To our knowledge, there is no study that has compared the Samsung Gear VR and Google Cardboard MR HMDs. However, other comparisons involving one of these two devices or other MR HMDs under 100 USD have already been carried out. For example, Buñ et al. [10] compared nVisor MH60V HMD, Oculus Rift and Samsung Gear VR for education using a virtual 3D human body atlas. A study involving 20 participants was carried out. In their study, features like the field of view, the weight of the devices and the general impressions and feelings of the participants were considered. They concluded that Oculus Rift and Samsung Gear VR could be effectively used for creating professional VR applications. Miller et al. [11] tested a VR application mimicking the game-day experience at Iowa State University. This application was displayed using a six-sided CAVETM, an Oculus Rift DK2 HMD and Merge VR goggles. A between-subjects user study ($N = 82$) compared the sense of presence between the different systems and a video control. The results revealed that while CAVETM scored the highest in presence, Oculus and Merge only experienced a slightly lower score compared to the CAVETM. Those authors suggested that Merge VR goggles is a viable alternative to CAVETM and Oculus.

If other types of HMDs (e.g., optical see-through) and comparisons (e.g., optical see-through vs. video see-through HMDs) are considered, different studies can be cited. For example, the work of Debernardis et al. [12] compared a monocular optical see-through HMD (Liteye LE 750A) with a video see-through HMD (Wrap 920AR Vuzix). They checked the effects on readability of these two HMDs using two backgrounds (light and dark), five colours (white, black, red, green and blue) and two text styles (plain text and billboarded text). The study involved 15 participants. They concluded that the readability turned out to be quicker on the optical see-through HMD. For the video see-through HMD, the background affected readability only in the case of text without billboard. They suggested that a good combination for indoor AR applications could be white text and blue billboard. In the work of Juan & Calatrava [13], a similar optical see-through HMD (Liteye 500) was compared with a video see-through HMD (5DT). They compared the sense of presence and anxiety in a non-phobic population (24 participants). If all of the participants were considered, they concluded that the system using the video see-through HMD induced a greater sense of presence than when the system was used with the optical see-through HMD. For the anxiety level, the system using the two HMDs provoked similar anxiety.

Samsung Gear VR and Google Cardboard have been used as VR HMDs, but, as mentioned above, it is also possible to use them for AR. Mostly, Perla & Hebbalaguppe [14] carried out a study to determine the feasibility of Google Cardboard for AR. Twenty engineers and research staff from an industry research lab participated in the study. The participants' task was to inspect a desktop computer using Google Cardboard. The interaction was by voice. The application was tested in two different settings: (A) a quiet reading room and (B) a work place (not too quiet and not too loud). The results indicated higher ratings for ease and usefulness of the application in the A setting (the quiet reading room). They argued that Google Cardboard could be a good solution for VR applications and a few other short-duration indoor AR applications and they discouraged these MR HMDs for outdoor activity applications.

This paper is not a comparison of the hardware of MR HMDs. This work focuses on the user's experience comparing a low-end MR HMD with a high-end MR HMD.

3. App Development

The AR app developed for the study used the same mobile phone (Samsung Galaxy Note 4) with a screen resolution of 1980×1080 pixels. The app was developed with the aim of obtaining user experience when using either of the two MR HMDs. Differences depended only on the MR HMD used.

Our app shows different 3D foods on a dish (one food at a time). Four different kind of foods are shown. The user can change from one food to another by simply placing his/her hand on a drip mat that acts as a virtual button. Once the 3D food is shown on the dish, the user can take the dish and look at it from different angles, try to touch the food, etc. (see Figure 2). The appearance and functionality of the app were exactly the same for both MR HMDs.



Figure 2. Participants using the app: (a) with Samsung Gear VR; (b) with Google Cardboard.

The reason why we chose the Samsung Galaxy Note 4 as the mobile device was its compatibility with both MR HMDs. The base application was developed using Unity3D (<http://unity3d.com>) and Vuforia SDK (<https://www.vuforia.com>). This base application had full functionality to run on Android devices. Unity 3D is a cross-platform game engine. It supports scripts written in C#. Vuforia SDK has an extension for Unity 3D. It facilitates the inclusion of animations and virtual objects. These inclusions are not as easy as if only Vuforia SDK is used. Vuforia uses computer vision techniques to recognize and track the following types of fiducial elements in real time: image targets, frame markers, multi-image targets, cylinder targets, virtual buttons, word targets and VuMark. Vuforia also includes recognition of 3D objects and physical environments (smart terrain) and provides support for HMDs such as Epson BT-200, ODG R-7, Samsung Gear VR and Microsoft HoloLens. Unity and Vuforia support the development of applications for Android and iOS devices. Our app runs on mobile devices with an Android Operating System. However, once the application has been developed, the platform for which the application has been built can be selected and the same code can be used for building applications for Android or iOS devices.

Other physical elements are used to superimpose the virtual elements: a dish and a drip mat. The dish is used to place the foods (see Figure 3a) and the drip mat (see Figure 3b) serves as a control mechanism (a virtual button that allows the object being displayed on the dish to be changed).



Figure 3. Physical elements used for the AR app: (a) Dish with the image target; (b) Drip mat with the image target.

The physical elements used to show the foods and to serve as a control mechanism have an image target superimposed on them. The images were previously processed with Photoshop. The Vuforia Target Manager (<https://developer.vuforia.com/target-manager>) was used to create the image targets (Figure 4).

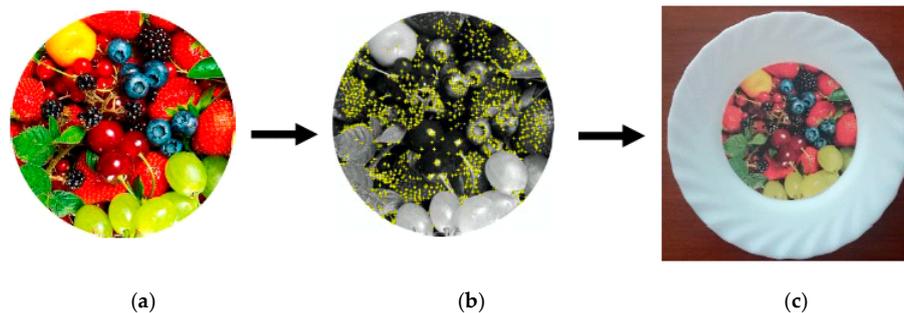


Figure 4. Image target: (a) Image target; (b) Features identified by Vuforia Target Manager; (c) Real dish with the image target.

The models used in the app were obtained from mainly two websites: www.blendswap.com and www.turbosquid.com. The models downloaded from Blend Swap were free, while the models downloaded from Turbo Squid required payment. After obtaining the models, it was necessary to perform modifications. We created several of the models. For the modification or creation of models, we used Blender (<https://www.blender.org/>), which is a free and open software for creating 3D models. Figure 5 shows two foods used in our app and how they were placed on the image target in the Unity scene.

The base application was adapted to work for Samsung Gear VR and Google Cardboard by configuring our Vuforia Unity scene for stereo displays and, thus, to support video-see-through digital eyewear apps. The real world captured by the camera is non-stereo, i.e., the two eyes perceive the same image. However, a different view of the virtual objects is perceived by each eye. Therefore, the user has 3D perception of the virtual objects. The perception is the same since a user can be wearing a binocular, video see-through HMD for AR and add only one camera to capture the real world. A detailed description of how to develop a Vuforia app for video see-through digital eyewear can be found in the Vuforia Developer Portal (<https://library.vuforia.com/articles/Solution/Working-with-Digital-Eyewear-in-Unity>). Figure 6 shows a screen capture of the app viewed with Google Cardboard.

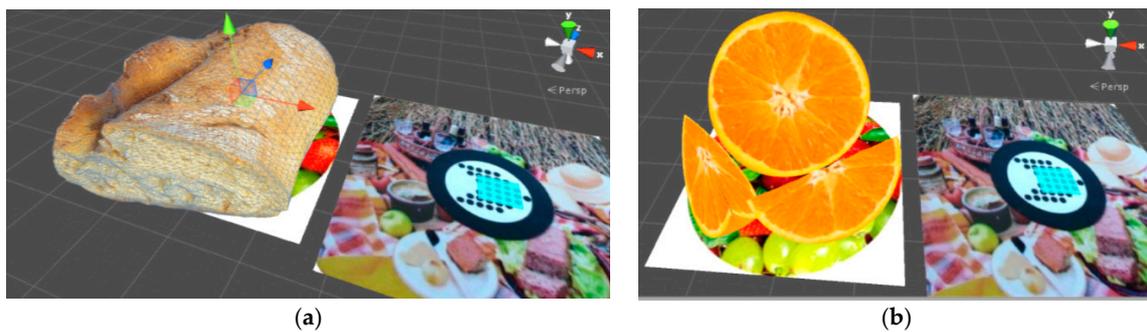


Figure 5. Two Unity scenes with two foods used in our app. (a) A piece of bread; (b) An orange.



Figure 6. Screen capture of the app viewed with Google Cardboard.

4. Study

Two previous questions were asked before enrolling the participants: (1) Do you have strabismus? (2) Do you not perceive 3D in 3D movies? Two users answered affirmatively to one of these questions and were discarded from the study. After of these two subjects, a total of 40 adults participated in the study. There were 20 men (50%) and 20 women (50%). They were between 19 and 58 years old. The mean age was 26.52 ± 7.57 years old. The total sample was divided into two groups. The participants were randomly assigned to each group, with the number of men and women being evenly distributed. In Group A, there were 10 men (50%) and 10 women (50%). In Group B, there were 10 men (50%) and 10 women (50%).

All of the participants were duly informed about the purpose of the study. The study was conducted according to the principles stated in the Declaration of Helsinki. The Ethics Committee of our university approved this study.

4.1. Measurements

A test and three different questionnaires were used to retrieve data for later analysis (see Figure 7):

- The Lang 1 Stereotest (Q1): The Lang 1 Stereotest has three objects. We followed the protocol suggested by Brown et al. [15]. A participant passed the test when he/she had 3/3 positive responses, 3/3 partial positive responses, or 2/3 positive and/or partial positive responses where the negative response was at the 550'' level. A participant failed the test when he/she had 3/3 negative responses and 2/3 negative responses where the single positive or partial positive response was at the 1200'' level. This test was applied to confirm that the participants had stereopsis.
- Users' perceptions about Samsung Gear VR (the Q2 questionnaire): This questionnaire was filled out by the user after using Samsung Gear VR. It consists of 26 questions, in which the user had to

evaluate each one on a Likert scale ranging from 1 “Totally disagree” to 7 “Totally agree”. The last two questions (OF#2 and #OF3) had different types of answers but they also used a 7 point Likert scale. The possible answers for question OF#2 ranged from 1 “Very poor” to 7 “Very good”. The possible answers for question OF#3 ranged from 1 “Much worse” to 7 “Much better” (see Table 2).

- Users’ perceptions about Google Cardboard (the Q3 questionnaire): The user filled out this questionnaire after using Google Cardboard. It contained the same questions as the previous questionnaire (Q2). Q3 had the same rating scale (see Table 2).
- MR HMD comparison (the Q4 questionnaire): Q4 was the last questionnaire to be filled out by the user regardless of the order of testing the devices. The user indicated with which device he/she had the best 3D experience, i.e., which one was more comfortable, which one he/she liked the most, which one he/she would like to buy and some other free questions (see Table 3).

Table 2. The results (median \pm interquartile range, M \pm IQ) for the Q2 and Q3 questionnaires. All of the questions used a 7 point Likert scale.

Id	Question	Q2 M \pm IQ		Q3 M \pm IQ	
CF#1	The application responded appropriately to my actions.	6 \pm 2	MA	6 \pm 1	MA
CF#2	The handling of the application was natural.	6 \pm 2	MA	6 \pm 2	MA
CF#3	The handling of the application was simple and uncomplicated.	7 \pm 1	SA	6.5 \pm 1	MA
CF#4	I did not notice delays between my actions and the expected results.	6 \pm 3.25	MA	6 \pm 1	MA
CF#5	I quickly got used to the application.	6 \pm 1	MA	6 \pm 1	MA
CF#6	The application was easy to control.	7 \pm 1	SA	7 \pm 1	SA
CF#7	At the end of the experience, I felt like an expert in handling the application.	6 \pm 1	MA	6 \pm 1	MA
DF#1	The control mechanisms did not distract me.	7 \pm 1	MA	6 \pm 1.25	MA
EF#1	The use of the MR HMD was comfortable.	6 \pm 2	MA	5 \pm 3	MA
EF#2	At no time did I think that I was going to drop the MR HMD	6.5 \pm 1	MA	5 \pm 3	MA
EF#3	The use of the MR HMD did not require a great effort from my arms.	7 \pm 1	MA	6 \pm 3	MA
EF#4	I did not feel any discomfort (dizziness, etc.) during the experience.	6 \pm 2.25	MA	6 \pm 2	MA
RF#1	I had the impression of seeing a dish with food.	6 \pm 1	MA	7 \pm 1	MA
RF#2	I had the impression that the feeds were real.	6 \pm 2	MA	6 \pm 1.25	MA
	It seemed to me that the food on the dish could be real food.				
RF#3	The food looks as real as the dish.	6 \pm 2	MA	6 \pm 1	MA
RF#4	I had the impression that the food was part of the real scene.	6 \pm 1.25	MA	6 \pm 1.25	MA
RF#5	I had the impression that I could touch and grab the food.	6 \pm 2	MA	6 \pm 2	MA
RF#6	I had the impression of seeing the foods as objects in 3D.	6 \pm 1	MA	6 \pm 1	MA
RF#7	I could examine the food closely.	7 \pm 1	SA	7 \pm 1	SA
RF#8	I could examine the food from different points of view.	7 \pm 1	SA	7 \pm 1	SA
RF#9	I did not pay attention to differences between the food and the dish.	6 \pm 2	MA	5 \pm 2	MA
RF#10	I did not have to make an effort to recognize the foods as 3D objects.	7 \pm 1	SA	7 \pm 1	SA
RF#11	I liked what the food looked like.	7 \pm 1	SA	6 \pm 1	MA
OF#1	I was involved during the experience.	6 \pm 1	MA	6.5 \pm 1	MA
OF#2	Score the 3D.	6 \pm 1	MA	6 \pm 1.25	MA
OF#3	This experience compared to other previous 3D experiences is...	5 \pm 2	MB	5 \pm 2	MB

Results for the Q2 and Q3 questionnaires.

Table 3. Results for the Q4 questionnaire. The possible answers for the questions were: (1). Samsung Gear (SG); (2). Google Cardboard (GC).

Id	Question	M \pm IQ	Answer
Q#1	Which device has a better 3D experience?	1 \pm 1	SG
Q#2	Which device did you find most comfortable?	1 \pm 0	SG
Q#3	Which device did you like the most?	1 \pm 1	SG
Q#4	If you had to buy one of them, which would you buy? (The price of the Samsung Gear viewer is €100 + the Samsung mobile. The price of the Google Cardboard viewer is less than €20 + any mobile up to 6 inches)	2 \pm 1	GC

Results for the Q4 questionnaire.

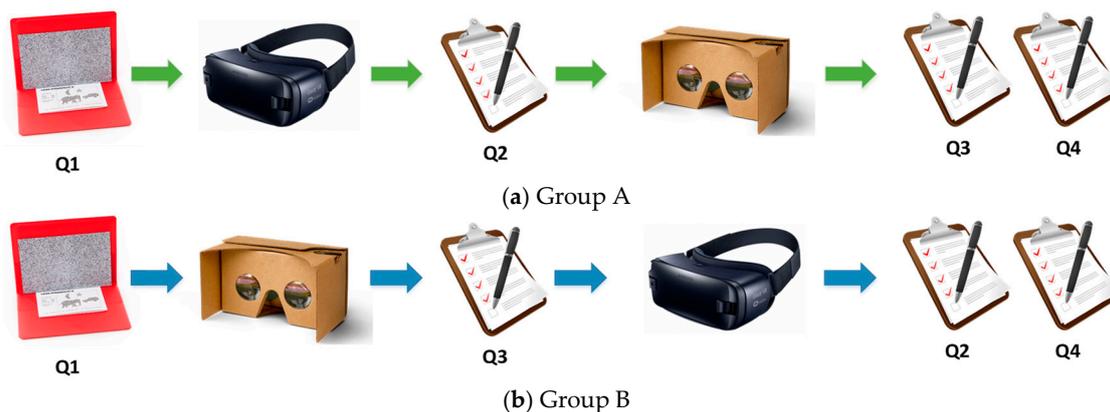


Figure 7. Procedure followed by Group A and Group B.

The Q2 and Q3 questionnaires measure the user's perceptions when using the app with a given MR HMD (Q2 with Samsung Gear VR and Q3 with Google Cardboard). The Q4 questionnaire collected the user's opinion about the comparison of the two MR HMDs.

Presence in virtual environments can be defined as an individual and context-dependent user's response that is related to the experience of "being there" [16]. According to Regenbrecht & Schubert [17], this definition cannot be applied exactly to AR. However, in AR, presence can also be achieved by measuring the presence of virtual elements into the real environment [17]. To measure this AR sense of presence, we added several questions to the Q2 and Q3 questionnaires. Two questionnaires, one from Witmer & Singer [18] and another one from Regenbrecht & Schubert [17] were adapted to our study. The Q2 questionnaire is detailed in Table 2. The factors considered are the following [18]:

- Control Factors (CF): degree of control, immediacy of control, anticipation of events, mode of control, physical environment modifiability.
- Sensory Factors (SF): sensory modality, environmental richness, degree of movement perception, active search.
- Distraction Factors (DF): selective attention, interface awareness.
- Realism Factors (RF): scene realism, information consistent with the objective world, meaningfulness of experience.

We also added questions related to ergonomics, preferences and other factors:

- Ergonomic Factors (EF): device comfort, effort, etc.
- Other Factors (OF): score of the 3D experience, involvement during the experience, 3D experience comparison.

4.2. Procedure

The participants in the study were divided into two groups:

- Group A: the participants used Samsung Gear VR first and afterwards they used Google Cardboard.
- Group B: the participants used Google Cardboard first and afterwards they used Samsung Gear VR.

These two groups tested both MR HMDs in different order. The following protocol was used (see Figure 7):

1. The Lang 1 Stereotest was applied to determine if the participants had stereopsis (Q1).
2. The participants tested the app using one MR HMD (Samsung Gear VR or Google Cardboard). Then, they filled out the User's Perceptions Questionnaire.

- a. Group A filled out the Q2 questionnaire (Samsung Gear VR).
 - b. Group B filled out the Q3 questionnaire (Google Cardboard).
3. Afterwards, they used the other MR HMD that they had not used in step 2.
- a. Group A filled out the Q3 questionnaire (Google Cardboard).
 - b. Group B filled out the Q2 questionnaire (Samsung Gear VR).
4. Finally, they filled out the Q4 questionnaire.

5. Results

An initial descriptive analysis was carried out. Data normality analysis reported that the study data did not fit a normal distribution. For this reason, the tests used were non-parametric (the Mann-Whitney *U* test for unpaired data, the Wilcoxon signed-rank sum test for paired data and the Kruskal-Wallis test). The data from the study was analysed using the statistical open-source toolkit R (<http://www.r-project.org>). For all of the tests, a $p < 0.05$ determined significance (in this case, the data is marked with two asterisks “**”).

5.1. Users' Perceptions Outcomes (Q2 and Q3)

The median of all of the questions is high (see Table 2). It is between 6 (Moderately agree) and 7 (Strongly agree) for both MR HMDs, with the exception of OF#3 whose median is 5 (Moderately better). If the questions are grouped by factors, the answers are:

- Control Factor Outcomes (CF): All of the questions (CF#1 to CF#7) have a median between 6 (Moderately agree) and 7 (Strongly agree) for both MR HMDs.
- Realism Factor Outcomes (RF): All of the questions (RF#1 to RF#11) have a median between 6 (Moderately agree) and 7 (Strongly agree) for both MR HMDs.
- Distraction Factor Outcomes (DF): The mean of question DF#1 using Samsung Gear VR is 7 (Strongly agree) and 6 using Google Cardboard (Moderately agree).
- Ergonomic Factor Outcomes (EF): All of the questions (EF#1 to EF#4) have a median of 6 (Moderately agree) for both MR HMDs.
- Other Factor Outcomes (OF): All of the questions (OF#1 to OF#3) have a median between 6 (Moderately agree) and 7 (Strongly agree) for both MR HMDs.

These results indicate that the participants were very satisfied with all of the app factors regardless of the MR HMD used.

The questions were analysed individually using a Wilcoxon signed-rank test for paired data. The analysis compares each individual question in Q2 with the same question in Q3. The questions that showed statistically significant differences were:

- EF#1 ($W = 509$, $Z = 3.594$, $p = 0.001$ **, $r = 0.402$), EF#2 ($W = 523$, $Z = 4.502$, $p = 0.001$ **, $r = 0.503$) and EF#3 ($W = 220$, $Z = 3.108$, $p = 0.002$ **, $r = 0.347$), all of them in favour of the Samsung Gear.
- RF#7 ($W = 74.5$, $Z = 2.008$, $p = 0.036$ **, $r = 0.225$) in favour of the Samsung Gear.

Specifically, OF questions evaluate the user's 3D experience. Question OF#1 shows that the users felt equally involved during the experience with both MR HMDs. The score of the 3D (OF#2) was similar for both MR HMDs. The users equally rated the 3D experience (OF#3) (compared to other experiences) independently of the MR HMD used.

5.1.1. Between-Subject Analysis

The users' perceptions were compared when the participants had only used one MR HMD (Group A used Samsung Gear VR and Group B used Google Cardboard). The data were analysed using a Mann-Whitney *U* test. There was a significant difference for the global score, considering all of the

questions ($U = 288$, $Z = 2.382$, $p = 0.018$ **, $r = 0.377$). The participants of Group A were more satisfied than the participants of Group B. The analysis was repeated to test if gender affected the users' scores. The analysis shows that gender did not influence the results (Group A: $U = 66.5$, $Z = 1.248$, $p = 0.226$, $r = 0.279$, Group B, $U = 71$, $Z = 1.589$, $p = 0.121$, $r = 0.355$). This implies that the participants who used Samsung Gear first were equally satisfied with the app regardless of gender. The same conclusion was obtained for those participants who used Google Cardboard. To determine if age affected the result, the unpaired Kruskal Wallis test was applied ($X^2(12) = 12.741$, $p = 0.388$, $r = 0.877$). This analysis indicated that the result was independent from age. The same conclusion was obtained for Google Cardboard ($X^2(11) = 14.559$, $p = 0.2036$, $r = 0.877$).

When the questions were classified by factors, there was only a statistically significant difference in the EF ($U = 307.5$, $Z = 2.925$, $p = 0.004$ **, $r = 0.462$) in favour of Samsung Gear VR. Age and gender did not affect the results when the questions were grouped.

5.1.2. Within-Subject Analysis

This subsection compares the users' perceptions of all of the participants, regardless of the order in which they used the MR HMDs (i.e., independently of the group they belonged to). Therefore, the study compares the scores of the Q2 and Q3 questionnaires. The results of the Wilcoxon signed-rank test show a statistically significant difference between the two groups in favour of Samsung Gear VR ($U = 1009.5$, $Z = 2.017$, $p = 0.044$ **, $r = 0.225$). Gender and age did not affect the result. To determine if the order of use of the MR HMDs (Group A or B) affected the results, an unpaired Kruskal-Wallis test was applied. The results indicated independence from the group factor.

The questions were also analysed grouping by factors. A Wilcoxon signed-rank test for paired data was applied to each group. The results indicated that there were no statistically significant differences for the CF, RF, DF and OF factors. However, the analysis shows that there was a statistically significant difference for the EF factor ($W = 583$, $Z = 4.386$, $p < 0.001$ **, $r = 0.49$) in favour of Samsung Gear VR. The tests determined that gender did not affect the users' scores for the two MR HMDs. However, there was an exception: gender affected the score for EF when the users tested Samsung Gear VR ($U = 272.5$, $Z = 1.973$, $p = 0.048$ **, $r = 0.312$) in favour of women. The unpaired Kruskal-Wallis test determines independence from age for all of the factors. The same conclusion was obtained for Google Cardboard. To determine if the order of use of the MR HMDs (Group A or B) affected the score, an unpaired Kruskal-Wallis test was applied. The analysis indicates independence from the group for all of the factors.

In another analysis for Group A, the scores for the Q2 questionnaire were compared with the scores for the Q3 questionnaire. The results of the Wilcoxon signed-rank test showed that there was a statistically significant difference in favour of Samsung Gear VR ($U = 275.5$, $Z = 2.043$, $p = 0.042$ **, $r = 0.323$). However, the gender and the age factors did not affect the results.

When analysing Group A scores grouped by factor, there were only significant differences in the EF ($U = 294$, $Z = 2.561$, $p = 0.011$ **, $r = 0.405$) in favour of Samsung Gear VR. Gender and age did not affect the users' scores when the questions were grouped by factors.

For Group B, the Wilcoxon signed-rank test shows that there was no statistical difference between the global score given by the users to the MR HMDs ($U = 233.5$, $Z = 0.907$, $p = 0.372$, $r = 0.143$). The results were independent from the gender and the age factors.

When analysing Group B scores grouped by factor, there was only a statistically significant difference for the EF ($U = 294$, $Z = 2.554$, $p = 0.011$ **, $r = 0.404$) in favour of Samsung Gear VR. The only factor affected by gender was the RF but only for Samsung Gear VR ($U = 80$, $Z = 2.274$, $p = 0.025$ **, $r = 0.508$). The women considered the experience to be more realistic with Samsung Gear VR than the men. Age did not affect the users' scores grouped by factors.

Figure 8 shows the total score of the Q2 questionnaire (Samsung Gear VR) and the Q3 questionnaire (Google Cardboard). The Samsung Gear VR total score was higher than the Google

Cardboard total score for both groups. The participants who rated the MR HMD the highest were the users in Group A when they used Samsung Gear VR.

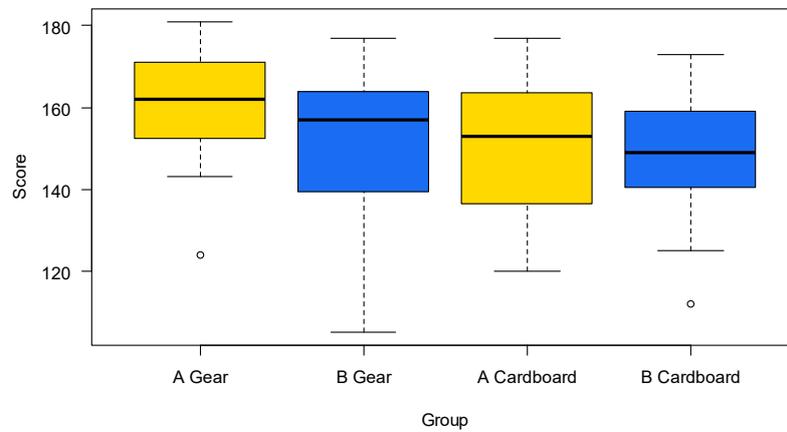


Figure 8. Total score by group.

The Spearman correlation (ρ) was used to test the relationship between factors. For the Samsung Gear VR, the results show a correlation from medium to high for all of the factors. The factors with significant correlations are shown in Figure 9.

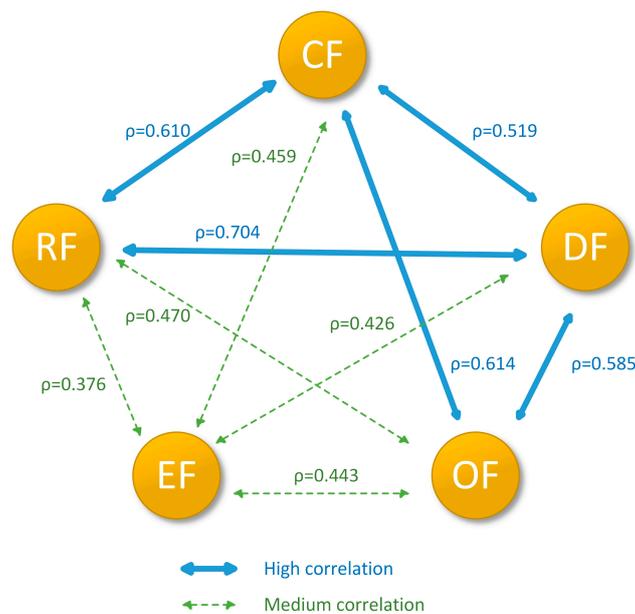


Figure 9. Samsung Gear VR. Factors with significant correlations.

For Google Cardboard, the Spearman tests showed lower correlation values. The factors with significant correlations are shown in Figure 10.

For Samsung Gear VR, there are 10 high/medium correlations. For Google Cardboard, there are 7 high/medium correlations. Therefore, the users of Samsung Gear VR estimated a closer association between the different factors considered. The correlation found between the EF factor and the DF factor for Samsung Gear VR and the non-existence of this correlation for Google Cardboard demonstrates the differences that the users expressed about the ergonomic problems.

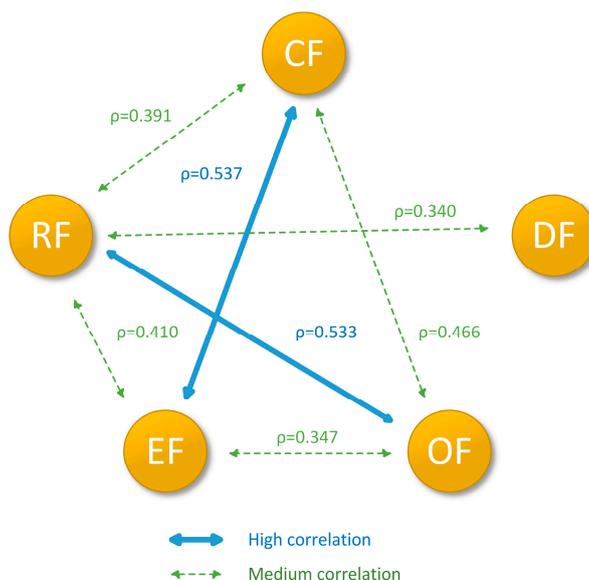


Figure 10. Google Cardboard. Factors with significant correlations.

5.2. Preferences about the MR HMD (Q4)

The Q4 questionnaire identifies the preference of the participants for Samsung Gear VR or for Google Cardboard. Table 3 shows this questionnaire and the median and the interquartile range of the scores for each question.

Figure 11 shows the number of users that chose each MR HMD for the Q4 questionnaire. The participants thought that Samsung Gear VR had better 3D (Q#1), it was more comfortable (Q#2) and they liked it more (Q#3). However, they would like to buy Google Cardboard (Q#4). The participants argued that the difference in quality does not justify the difference in price.

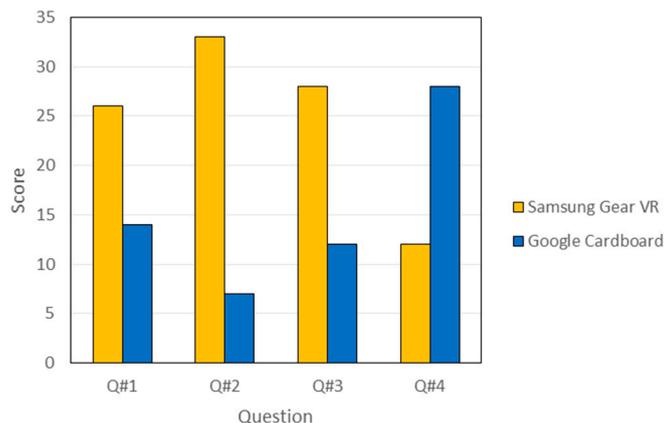


Figure 11. The Q4 questionnaire. Number of users.

The score variable combines all of the answers of the questions related to preferences for a device (Q#1 to Q#4). To determine if gender affected the score, an unpaired Mann-Whitney *U* test was performed ($U = 180.5, Z = -0.543, p = 0.597, r = 0.086$). The results showed that there were no statistically significant differences for the global score between men and women. A Kruskal-Wallis test was performed to determine whether or not the age affected the global score ($X^2(15) = 19.871, p = 0.1769, r = -0.8662585$). The results indicated that age does not affect it. To determine if the group factor (A or B) affected the score, an unpaired Mann-Whitney *U* test was performed ($U = 142,$

$Z = -1.614$, $p = 0.110$, $r = 0.255$). The results showed that the order in which the participants used the devices does not affect the result.

The Q4 questionnaire contains some free answer questions (see Table 4). Question F#1 asks the user why they would buy one MR HMD or another. The participants that would like to buy Samsung Gear VR argued that this MR HMD offers a better 3D experience, the quality is higher and the ergonomic factor is better. However, participants that would like to buy Google Cardboard argued that the relationship between quality and price is better. Besides, some participants added that design of Google Cardboard is simple and nice. Some answers to question F#2 about Samsung Gear VR indicated that it is more comfortable and has higher quality than Google Cardboard. Some answers to question F#3 about Google Cardboard indicated that it is cheaper, it has good quality for the price, it is simple and it has an original design. However, they added that the comfort must be improved. Some answers to F#4 indicated that what they liked the most were the realism of the foods or the immersion experience.

Table 4. Free answer questions of the MR HMD comparison questionnaire (Q4).

Id	Question
F#1	Why? (About Q#4)
F#2	What do you think about Samsung Gear VR?
F#3	What do you think about Google Cardboard?
F#4	What did you like most about the experience?

6. Discussion

This paper presents a comparative study of two MR HMDs (low-end and high-end). The most notable differences between low-end and high-end MR HMDs are: high-end MR HMDs usually have higher visual quality and can be adjusted to the user's head; low-end MR HMDs are cheaper. Samsung Gear VR was chosen to represent high-end MR HMDs and Google Cardboard was chosen to represent low-end MR HMDs since they are two of the most popular MR HMDs on the market.

This work attempted to determine if the technical differences between the two MR HMDs really create a significant difference in the user's experience. Our study involved 40 participants. We developed an AR app to evaluate the visual quality of the user's experience using the two MR HMDs. The users' perceptions depended only on the MR HMD used. The results showed that there were no statistically significant differences in the users' perceptions, which corroborates our first hypothesis. The scores were high for both MR HMDs. From the results, we can argue that the users had a satisfactory experience using the two MR HMDs. Our positive results are in line with other previous positive experiences using MR HMDs (not only the two MR HMDs used in our study but also the Merge VR goggles) [7,8,10,11].

The users' perceptions were not dependent on gender, age or the order in which the participants used the MR HMDs (Group A or B). The analysis of the factors evaluated by the questionnaires showed that there were no statistically significant differences between the two MR HMDs. These results suggest that regardless of gender or age, everyone can have rich AR experiences with MR HMDs of this type.

With regard to the stereo visualization using a monocular camera for capturing the real world, the participants did not mention any problem in this regard. The person in charge of the validations did not observe that the participants had any problems in this sense. Our work is in line with previous works in which the visualization was stereo and the real world was captured using a monocular camera, using video-see through HMDs (e.g., attaching a monocular camera to an Oculus Rift HMD), as well as using other combinations (e.g., combining an autostereoscopic display with a monocular camera [19]).

With regard to the ergonomics factor, there was a statistically significant difference in favour of the high-end MR HMD, which corroborates our second hypothesis. Moreover, if the correlations are considered, the correlation found between the EF factor and the DF factor for the high-end MR HMD and the non-existence of this correlation for the low-end MR HMD also demonstrates this result.

Our explanation for this is that the user has to hold the low-end MR HMD using his/her hand and at the same time pick up and move the dish to observe the different foods as well as to put his/her hand on the drip mat to change the food. In other words, the fact that the low-end MR HMD does not have elements for its adjustment to the user's head has mainly influenced this result. This argument was corroborated by some comments of the participants and the observations made by the person in charge of the study. However, another study could be carried out to determine the extent to which this and other factors (e.g., the difference in the field of view, the difference of sensors or the weight) influence ergonomic perceptions. Our recommendation is to use MR HMDs with a head restraint when the user has to manipulate real objects. This problem is similar to the problem of autostereoscopic mobile devices for AR. In that case, the user has to hold the device with his/her hands and focus his/her eyes on an area of the screen in order to not lose depth perception. Meanwhile, he/she has to interact with real objects or even touch an area of the screen, making it very difficult to maintain depth perception. This problem is less severe but it also occurs in mobile AR applications since the users have to hold the mobile device in their hands while they perform another activity (e.g., drawing). For example, in the study carried out by Gombač et al. [20], the participants held a mobile device in their hands while drawing and their performance drastically decreased.

When the users were asked about their preferred MR HMD, the participants mostly chose the high-end MR HMD. When they were asked about which device they would buy, the participants mostly responded they would buy the low-end MR HMD. In low-end MR HMDs, apart from the lenses, the rest of the MR HMD is made of cardboard. Users can even download the parts and assemble them by themselves. This makes it even cheaper. In addition, the material with which it is built makes it fully customizable and recyclable. The participants may have identified these possibilities as advantages, in addition to its lower price. Our argument is that these aspects greatly influenced their choice.

According to Boos et al. [3], some of the drawbacks of current MR HMDs are the following: low graphical quality, poor battery life and uncomfortable thermal radiation. Mobile GPU rendering can consume up to 20 W of peak power [21] and therefore, the thermal output is a safety concern for near-eye devices without active cooling. These three drawbacks should be improved so that MR HMDs can be increasingly comparable to tethered HMDs.

In our study, the participants were recruited from our university. The study was carried out in controlled environments. However, a study conducted "in the wild" could be carried out in the future. The term "in the wild" has the connotation of performing studies with real users in uncontrolled environments rather than in laboratories [22]. Larger samples could be collected using this mode of data collection. Steed et al. [9] demonstrated the feasibility of running studies on VR using MR HMDs in the wild.

We used Vuforia for the development of our app. However, with the presentation of the ARKit (<https://developer.apple.com/arkit>) and Google ARCore (<https://developers.google.com/ar>) frameworks (middle of 2017), new alternatives have appeared. ARKit and ARCore are two new frameworks for building AR apps. ARKit is for iPhone and iPad and ARCore is for Android devices. ARKit and ARCore can create AR experiences without the need for placing additional elements to the real scene (e.g., markers or image targets). This trend was initiated by the Tango SDK (<https://developers.google.com/tango>). The problem with Tango SDK is that it can only be used for building AR apps for Android devices with special features (e.g., Lenovo Phab 2 Pro). However, ARCore works on Google's Pixel twins and Samsung's Galaxy S8 running either Android Nougat or Android Oreo. It has been announced that ARCore will support all devices running Android 7.0 Nougat and above.

7. Conclusions

We have compared two different MR HMDs (low-end and high-end). To our knowledge, this is the first comparison involving those two different MR HMDs. The results showed that there were no statistically significant differences in the user's perceptions between the two MR HMDs. Only

the ergonomic factor offered a statistically significant difference in favour of the high-end MR HMD. However, when the users were asked about which device they would buy, the participants mostly responded they would buy the low-end MR HMD. Our argument is that the price and the possibility of building it yourself greatly influenced this result. Therefore, from our findings, we can argue that the two MR HMDs and other similar MR HMDs could be used to have rich AR experiences at an affordable cost.

In this paper, we have compared two MR HMDs. However, several other comparisons are possible for future work (e.g., other MR HMDs or comparing MR HMDs with tethered HMDs). The sense of presence induced by MR HMDs and the cyber sickness that can be induced by these MR HMDs are other aspects to be studied. Also, new applications for different purposes should be developed to be visualized with current MR HMDs. We hope to benefit users and researchers of MR HMDs with the work and ideas presented here.

Acknowledgments: We would like to thank the users who participated in the study for their contributions. We also thank the reviewers for their valuable suggestions.

Author Contributions: M.-Carmen Juan conceived and designed the study; Richard López developed the apps. Richard López, Inmaculada García and Ramón Mollá performed the experiments; Richard López, Inmaculada García and M.-Carmen Juan analyzed the data; M.-Carmen Juan, Inmaculada García and Ramón Mollá wrote the paper.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Bowman, D.; Kruijff, E.; Laviola, J.; Poupyrev, I. *User Interfaces: Theory and Practice*; Addison-Wesley: Boston, MA, USA, 2004.
2. Usoh, M.; Slater, M. An exploration of immersive virtual environments. *Endeavour* **1995**, *19*, 34–38. [[CrossRef](#)]
3. Boos, K.; Chu, D.; Cuervo, E. FLASHBACK: Immersive Virtual Reality on Mobile Devices via Rendering Memoization. *GetMobile* **2016**, *20*, 23–27. [[CrossRef](#)]
4. Markets and Markets Report. Augmented Reality Software Market—Global Forecast to 2022. 2007. Available online: <http://www.researchandmarkets.com/reports/4228225/augmented-reality-software-market-software#pos-0> (accessed on 9 October 2017).
5. Markets and Markets Report. Mobile Augmented Reality Market by Component, Application, Vertical and Geography—Global Forecast to 2022. 2016. Available online: <http://www.marketsandmarkets.com/> (accessed on 10 January 2018).
6. Kesselman, M. Current CITE-ings from the popular and trade computing literature: Google Cardboard—Virtual reality for everyone. *Libr. Hi Tech. News* **2016**, *33*, 15–16. [[CrossRef](#)]
7. Mosadeghi, S.; Reid, M.W.; Martinez, B.; Rosen, T.; Spiegel, B.M.R. Feasibility of an Immersive Virtual Reality Intervention for Hospitalized Patients: An Observational Cohort Study. *JMIR Ment. Health* **2016**, *3*, E28. [[CrossRef](#)] [[PubMed](#)]
8. Tashjian, V.C.; Mosadeghi, S.; Howard, A.R.; Lopez, M.; Dupuy, T.; Reid, M.; Martinez, B.; Ahmed, S.; Dailey, F.; Robbins, K.; et al. Virtual Reality for Management of Pain in Hospitalized Patients: Results of a Controlled Trial. *JMIR Ment. Health* **2017**, *4*, E9. [[CrossRef](#)] [[PubMed](#)]
9. Steed, A.; Friston, S.; Murcia López, M.; Drummond, J.; Pan, Y.; Swapp, D. An ‘In the Wild’ Experiment on Presence and Embodiment Using Consumer Virtual Reality Equipment. *IEEE Trans. Vis. Comput. Graph.* **2016**, *22*, 1406–1414. [[CrossRef](#)] [[PubMed](#)]
10. Buñ, P.; Górski, F.; Wichniarek, R.; Kuczko, W.; Hamrol, A.; Zawadzki, P. Application of Professional and Low-cost Head Mounted Devices in Immersive Educational Application. In Proceedings of the International Conference on Virtual and Augmented Reality in Education, Monterrey, Mexico, 9–21 November 2015; pp. 173–181. [[CrossRef](#)]
11. Miller, J.; Baiotto, H.; MacAllister, A.; Hoover, M.; Evans, G.; Schlueter, J.; Kalivarapu, V.; Winer, E. Comparison of a Virtual Game-Day Experience on Varying. *Electron. Imaging* **2017**, 30–37. [[CrossRef](#)]

12. Debernardis, S.; Fiorentino, M.; Gattullo, M.; Monno, G.; Uva, A.E. Text Readability in Head-Worn Displays: Color and Style Optimization in Video versus Optical See-Through Devices. *IEEE Trans. Vis. Comput. Graph.* **2014**, *20*, 125–139. [[CrossRef](#)] [[PubMed](#)]
13. Juan, M.C.; Calatrava, J. An Augmented Reality System for the Treatment of Phobia to Small Animals Viewed Via an Optical See-Through HMD: Comparison With a Similar System Viewed via a Video See-Through HMD. *Int. J. Hum. Comput. Interact.* **2011**, *27*, 436–449. [[CrossRef](#)]
14. Perla, R.; Hebbalaguppe, R. Google Cardboard Dates Augmented Reality: Issues, Challenges and Future Opportunities. *arXiv* **2017**, arXiv:1706.03851.
15. Brown, S.; Weih, L.; Mukesh, N.; McCarty, C.; Taylor, H. Assessment of adult stereopsis using the Lang 1 Stereotest: A pilot study. *Binocul. Vis. Strabismus Q.* **2001**, *16*, 91–98. [[PubMed](#)]
16. Bowman, D.; McMahan, R. Virtual Reality: How much immersion is enough? *Computer* **2007**, *7*, 36–43. [[CrossRef](#)]
17. Regenbrecht, H.; Schubert, T. Measuring Presence in Augmented Reality Environments: Design and a First Test of a Questionnaire. In Proceedings of the Fifth Annual International Workshop Presence, Porto, Portugal, 9–11 October 2002.
18. Witmer, B.; Singer, M. Measuring Presence in Virtual Environments: A Presence Questionnaire. *Presence* **1998**, *7*, 225–240. [[CrossRef](#)]
19. Arino, J.J.; Juan, M.C.; Gil-Gómez, J.A.; Mollá, R. A comparative study using an autostereoscopic display with augmented and virtual reality. *Behav. Inf. Technol.* **2014**, *33*, 646–655. [[CrossRef](#)]
20. Gombač, L.; Čopič Pucihar, K.; Kljun, M.; Coulton, P.; Grbac, J. 3D Virtual Tracing and Depth Perception Problem on Mobile AR. In Proceedings of the CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '16), San Jose, CA, USA, 7–12 May 2016; pp. 1849–1856. [[CrossRef](#)]
21. Jin, T.; He, S.; Liu, Y. Towards accurate GPU power modeling for smartphones. In Proceedings of the 2nd Workshop on Mobile Gaming, MobiGames@MobiSys, Florence, Italy, 19 May 2015; pp. 7–11. [[CrossRef](#)]
22. Brown, B.; Reeves, S.; Sherwood, S. Into the wild: Challenges and opportunities for field trial methods. In Proceedings of the Special Interest Group on Computer-Human Interaction (SIGCHI) Conference on Human Factors in Computing Systems, Vancouver, BC, Canada, 7–12 May 2011; pp. 1657–1666. [[CrossRef](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).