



**HAL**  
open science

# Studying Exocentric Distance Perception in Optical See-Through Augmented Reality

Etienne Peillard, Ferran Argelaguet Sanz, Jean-Marie Normand, Anatole Lécuyer, Guillaume Moreau

► **To cite this version:**

Etienne Peillard, Ferran Argelaguet Sanz, Jean-Marie Normand, Anatole Lécuyer, Guillaume Moreau. Studying Exocentric Distance Perception in Optical See-Through Augmented Reality. ISMAR 2019 - 18th IEEE International Symposium on Mixed and Augmented Reality, Oct 2019, Beijing, China. pp.1-8. hal-02200822

**HAL Id: hal-02200822**

**<https://hal.science/hal-02200822v1>**

Submitted on 31 Jul 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Studying Exocentric Distance Perception in Optical See-Through Augmented Reality

Etienne Peillard\*<sup>†</sup> Ferran Argelaguet\* Jean-Marie Normand<sup>†</sup> Anatole Lécuyer\* Guillaume Moreau<sup>†</sup>

\*Inria, Univ. Rennes, IRISA, CNRS  
<sup>†</sup>École Centrale de Nantes, AAU, Inria Hybrid

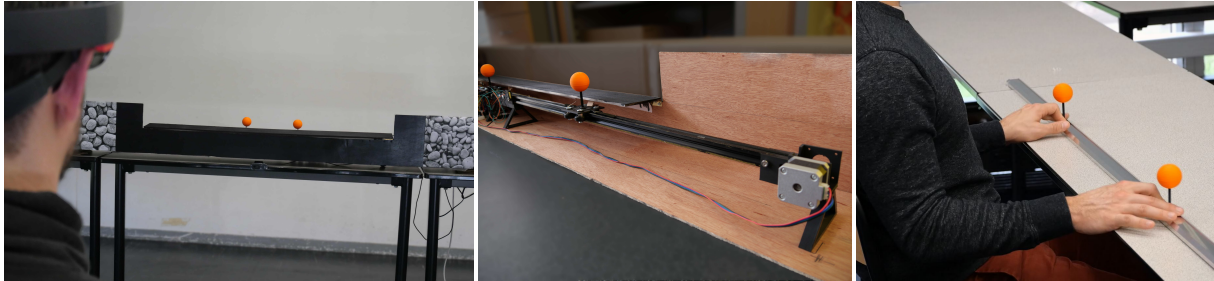


Figure 1: Left, bench displaying two real spheres. The hinge-actuated moving panel, opened here, could be automatically opened/closed to reveal/hide the visual stimuli. Center, one of the two rails of the bench, seen from behind. An orange sphere is attached on top of a trolley that can slide on the rail. The trolley is moved by a stepper motor through a belt. The other half of the bench is symmetrical. Right, participants could provide the perceived exocentric distance by placing two sliding spheres. After the participants placed the spheres the system automatically took a picture of both spheres which was used to measure the distance between both spheres.

## ABSTRACT

While perceptual biases have been widely investigated in Virtual Reality (VR), very few studies have considered the challenging environment of Optical See-through Augmented Reality (OST-AR). Moreover, regarding distance perception, existing works mainly focus on the assessment of egocentric distance perception, i.e. distance between the observer and a real or a virtual object. In this paper, we study exocentric distance perception in AR, hereby considered as the distance between two objects, none of them being directly linked to the user. We report a user study ( $n=29$ ) aiming at estimating distances between two objects lying in a frontoparallel plane at 2.1m from the observer (i.e. in the medium-field perceptual space). Four conditions were tested in our study: real objects on the left and on the right of the participant (called *real-real*), virtual objects on both sides (*virtual-virtual*), a real object on the left and a virtual one on the right (*real-virtual*) and finally a virtual object on the left and a real object on the right (*virtual-real*). Participants had to reproduce the distance between the objects by spreading two real identical objects presented in front of them. The main findings of this study are the overestimation (20%) of exocentric distances for all tested conditions. Surprisingly, the *real-real* condition was significantly more overestimated (by about 4%,  $p = .0166$ ) compared to the *virtual-virtual* condition, i.e. participants obtained better estimates of the exocentric distance for the *virtual-virtual* condition. Finally, for the *virtual-real/real-virtual* conditions, the analysis showed a non-symmetrical behavior, which suggests that the relationship between real and virtual objects with respect to the user might be affected by other external factors. Considered together, these unexpected results illustrate the need for additional experiments to better understand the perceptual phenomena involved in exocentric distance perception

\*surname.lastname@inria.fr

<sup>†</sup>surname.lastname@ec-nantes.fr

with real and virtual objects.

**Keywords:** Perception, Distance, Augmented Reality, User Experiment, Psychophysical Study

**Index Terms:** Human-centered computing—Human computer interaction (HCI)—Interaction paradigms—Virtual reality; Human-centered computing—Human computer interaction (HCI)—HCI design and evaluation methods—User studies

## 1 INTRODUCTION

Augmented Reality (AR) intends to fill the gap between real and digital content. However, when it comes to vision, display techniques are not yet able to make virtual objects indistinguishable from their real counterparts. As a consequence, observers have a different perception of virtual and real objects in AR and can thus be biased in many ways when comparing the perceived characteristics of these objects. In particular, the specific case of distance perception is a topic where major differences have been reported.

However, beyond augmented and virtual reality, regardless of the medium considered, space perception is a wide field of research. Many psychological studies tackled human space perception through depth perception, perceptual illusions which arise when combining distance, orientation and speed perception, and also objects parameters such as size or height. Distance perception can differ greatly according to the experimental apparatus and the task considered [20]. Overall, previous studies state that rendering parameters and environmental cues are two of the main factors which can bias the perception of digital objects [3, 14].

Interestingly and despite the recent development of AR devices, in particular Optical See-Through (OST) devices, very little work has been dedicated to characterizing perceptual biases related to distance perception between real and virtual objects displayed in AR. Moreover, most research was performed on egocentric distance perception, i.e. distance between the observer and objects of the environment, while several AR tasks (including discrepancy checks, maintenance or project reviews) need to evaluate inter-objects spatial relationships, i.e. exocentric distances. This paper presents, to

the best of our knowledge, the first study using a direct measurement technique in order to evaluate the potential bias of exocentric distance perception in AR.

The remainder of this paper is structured as follows: Sect. 2 is dedicated to describing relevant related work in the field of distance perception, being in real, virtual as well as augmented environments. Sect. 3 presents our user experiment dedicated to studying exocentric distance perception in OST AR. Finally, after discussing our results in Sect. 4, we conclude and present future research leads in Sect. 5.

## 2 RELATED WORK

Distance perception is generally divided into two categories: egocentric and “non-egocentric”, see [19]. Indeed, while it is easy and straightforward to define egocentric distance perception as evaluating a distance between an object and the observer, “non-egocentric” can cover various definitions: *exocentric distance* perception when it comes to evaluate the distance between two objects, *allocentric distance* perception when considering positions and distances in a non-egocentric frame of reference, and *objects characteristics* such as size estimation. This section will mainly focus on exocentric distance perception, in particular on results related to our experimental setup (see Sect. 3.4), namely exocentric distance perception in real, virtual and augmented environments.

### 2.1 Foreword on Distance Perception: Distance Field Theories and Reporting Techniques

Humans’ perceptual space is generally divided into three areas: personal space, action space, and vista space, see [30]. The personal space corresponds to distances within 1.5 meters (a little over arm’s length), where people can perceive depth accurately in reality. The action space lies between 1.5 and 30 meters, where people can talk to each other. In the action space, depth perception becomes underestimated for stationary observers. The vista space refers to distances over 30 meters, where distance underestimation and the distance between the object and the observer are directly correlated. In the following, we will try to classify existing experiments using Cutting’s classification but unfortunately, not all experiments lie in a single depth field.

Before reporting existing results, we also need to recall that there exist many different ways of reporting distances. The most representative reporting techniques are the following: (i) matching where the observer has to align a visible reference (e.g. a finger, etc.) with the virtual object in an open-loop procedure; (ii) reaching, where the observer has to align a non-visible reference (e.g. a finger while the hand is hidden behind a panel) to a virtual object; (iii) two-alternative forced choice (2AFC) where the observer has to choose between two proposed options; (iv) verbal reporting where the observer has to announce the estimated distance orally and (v) blind walking, where the observer has to cover the estimated distance between himself and the virtual object while blind walking. It should be noted that the first three possible reporting methods only work for short distances. Obviously the choice of reporting method has to be carefully made since it has an important impact on the results [28, 29].

### 2.2 Exocentric distance perception in real life

One of the first studies focusing on a comparison between egocentric and exocentric distance perception in real environments is that of Loomis et al. [19]. Authors used a matching task where users had to adjust two metallic rods placed on the observer’s frontal plane so that the distance between them matched that of two target metallic rods that were placed on the depth plane (i.e. the sagittal plane, being perpendicular to the frontal one). The distance between the rods lied within the 4-12m range. The question motivating this work was whether natural environments viewed binocularly from a stationary vantage point are perceived correctly. They found that there was an underestimation of egocentric distances compared to exocentric

distances. In [16], Kudoh reproduced the work from Loomis et al. [19] using blind walking as a reporting technique and confirmed the overall overestimation of exocentric distances.

An hypothesis presented in [19] and confirmed in [10, 21] is that there is a dissociation between the perception of location and shape. Indeed, while our perception of location is accurate (up to at least 20m under full cue conditions), the perception of 2D shapes viewed under the same conditions exhibits large systematic distortions. Given these results, this hypothesis further assumes that the perception of extent and of shape is based on different neural computations from those involved in the perception of location.

Other studies in real environments [9, 22] showed that stimuli lying in the depth (i.e. sagittal) plane are indeed underestimated compared to stimuli placed in the frontal plane. In a later work, Foley et al. [10] proposed a mathematical model for the relationship between perceived and physical objects (i.e. visual angles) as well as perceived distances. Nevertheless, their model showed that this relationship cannot be represented by any Euclidean or other metric geometry.

Levin and Haber [17] confirmed that a change in the viewing angle can modify the exocentric depth perception of exactly the same stimuli. Orientation has also been shown to be an issue [2, 17].

Finally, Predebon [24] investigated the effect of familiar and unfamiliar size of objects in the vista space (between 10m and 187.5m). Results showed that the familiar/unfamiliar characteristics of stimuli do not influence depth-matching judgments, judgments of relative depth, nor judgments of the relative lengths and relative egocentric distances of extents presented in the observers’ frontoparallel plane. Finally, it was shown that familiar sizes influence judgments of the relative egocentric distances of unfamiliar and familiar objects.

To summarize, the very extensive literature regarding exocentric distance perception in real environments seems to agree on the fact that exocentric distance of stimuli placed in the sagittal plane is underestimated compared to that of similarly distant objects positioned in the frontoparallel plane.

### 2.3 Exocentric distance perception in VR

Regarding exocentric distance perception in virtual environments, results are less systematic since they also depend on the type of Virtual Reality (VR) display used in the experiments.

In [31], Waller showed an overestimation of exocentric (i.e. inter objects) distances and that this systematic error based on verbal reporting could be improved when the experimenter gave participants direct feedback after their verbal report. The experimental conditions allowed the participant to move freely and two display conditions were tested: Head Mounted Display (HMD)-based and desktop-based VR. Nevertheless, the Virtual Environment (VE) was very basic (a huge gray cube with a regular grid superimposed on the cube’s faces) and thus not very representative of nowadays’ VE.

Conversely, Wartenberg and Wiborg [32] allowed participants to move but also observed distance overestimation in the VR condition (within a CAVE) while the desktop conditions lead to both over- and underestimations.

Regarding size estimation, there is no complete consensus: while some authors tend to demonstrate a certain accuracy [11, 15], [26] shows that virtual objects are estimated as smaller than in reality. It should be noted that Stefanucci et al. [26]’s setup used flat TV screens to display 3D objects in a “desktop VR” configuration unlike [11, 15] that were conducted in “immersive VR”.

In [12], Geuss et al. showed that, while wearing an HMD, exocentric distances are compressed when evaluated in the depth plane but are accurate when objects are located in the frontal plane. Based on Geuss et al. [12]’s findings, Kelly et al. [14] proposed an alternative hypothesis to explain these results: they were due to a combination of underestimation caused by the VE and overestimation of frontal relative to depth extents. They thus conducted two experiments in

which they reproduced Geuss et al.’s findings in a similar VE but also demonstrated an underestimation in both frontal and depth planes when using a simpler VE composed of a single textured plane.

In [18], Lin et al. asked participants to sketch a 2D line representing the center-to-center distance between a real object and a virtual object presented using a 3D stereoscopic projector onto a screen. They showed that participants underestimated the distance and that the accuracy worsened along with an increasing distance between both objects. They also showed that aligning both objects in the vertical direction reduced this underestimation.

Finally, still using a projection-based stereoscopic environment, Woldegiorgis et al. [33] showed that the accuracy of a pointing task when objects lie in the frontal plane was lower (85%) compared to that of the physical world (94%). They also showed that this accuracy was affected by the egocentric distance between the participant and the frontal plane.

To summarize, the findings in VR that are the most related to our present study show that exocentric distances in VEs visualized through an HMD are compressed when evaluated in the depth plane but are accurate when objects are located in the frontal plane.

## 2.4 Exocentric Distance Perception in AR

Finally, we conclude this section with an overview of the few studies dedicated to exocentric distance perception in AR which is the topic of our study. Dey et al. [5] showed that in hand-held AR devices, there was an underestimation (using verbal reporting) of exocentric distances and egocentric distances in the vista space (distances from 30m to 110m were considered in the paper). Nevertheless, exocentric distances were only evaluated between virtual objects as no real objects were involved.

Then, Sugano et al. [27] showed that the presence of virtual shadows increased participants’ accuracy in a vertical ordering task using a Video See-Through (VST) HMD setup.

Finally, Pointon et al. [23] studied space perception by asking participant if they could pass through a virtual aperture between two poles, presented at different widths and distances. Their results showed no difference between virtual poles in AR and real ones.

From the presented literature review, exocentric distance perception, and while having received substantial interest from the community both in real and virtual environments, exocentric distance perception has been relatively left aside in augmented environments. In particular, there is no study focusing on exocentric distance perception in OST AR environments. This is an important and relevant matter, especially given the recent development of high quality consumer grade OST AR headsets, such as the Microsoft HoloLens. Therefore, in this paper, we focused on the study of exocentric distance perception in different visual conditions in an OST AR environment, with real and/or virtual objects.

## 3 USER EXPERIMENT

In this section we present a user experiment aimed at the evaluation exocentric distance perception in OST AR. To do so, we designed an experiment where participants had to evaluate distances between objects (spheres placed in a plane perpendicular to the users’ view) in the following configurations: (i) two real spheres, (ii) two virtual spheres displayed in AR and (iii/iv) one real sphere and one virtual sphere. This last condition is in fact two-fold, since the real sphere can be displayed on the left hand side or on the right hand side of the participant. In the following, we detail the experimental design, the apparatus, as well as the protocol.

### 3.1 Participants

This experiment involved 29 participants (21 male and 8 female), aged from 19 to 54 years (mean=25.0, SD=7.2). Among them, 27 were right-handed, one was left-handed and one reported to be ambidextrous. The eye dominance was assessed using the method

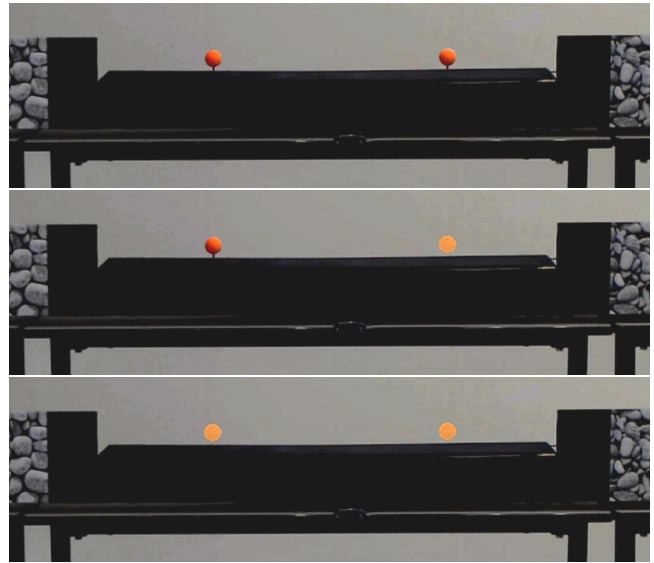


Figure 2: Three out of the four possible display conditions, from top to bottom: *real-real*, *real-virtual* (note that *virtual-real* is also part of the experimental design) and *virtual-virtual*.

proposed by Durand [6] in which participants were asked to look at a distant object through a hole made with their two hands. 9 participants had a dominant left eye and 20 had a dominant right eye. Participants were students of the university or members of the laboratory and naive about the purpose of the experiment. They were not paid for their participation. All participants had normal or corrected-to-normal vision and 23 out of 29 had already used an HMD before but none of them had used an OST AR HMD before.

### 3.2 Experimental Apparatus

The real stimuli were orange 3D-printed spheres with a radius of 20 mm. They were placed on a bench located 2.1 meters away from the participant (see Fig. 1 left). This bench was composed of two symmetric rails, one for each sphere. A moving trolley was placed on each rail and controlled by a stepper motor. Given the mechanical characteristics of the motor and of the transmission system, spheres could be positioned with a precision of 0.2 mm. The control software of the bench run on a Raspberry Pi and was written in Python.

The placing mechanism was completely hidden by a black wooden structure. In addition, a *moving panel* was cut inside and could be toggled on and off in order to reveal or hide the visual stimuli. Figure 1 (center) allows to better appreciate the mechanism in which the moving panel is opened and the two real spheres are visible. All the other parts of the bench were always hidden during the experiment. AR markers were placed at each end of the panel which were used to calibrate the position of the bench in the VE.

The virtual stimuli were spheres displayed on an OST AR headset, the Microsoft HoloLens. The HoloLens has a field of view (FOV) of  $30^\circ \times 17.5^\circ$  with a resolution of  $1268 \times 720$  px per eye. It contains an inertial measurement unit (IMU) and four sensors (a depth camera, a RGB camera, a four-microphone array, and an ambient light sensor) used to map the 3D environment in real-time. The HoloLens also supports voice input, basic gesture recognition, and head tracking. We selected the HoloLens mainly due to the quality and robustness of its tracking. The AR scenes were built using Unity3D v2018.3.0f2. The spheres were displayed using the standard Unity shader, they were lit with a single frontal directional light and without texture. This provided a coherent ambient and specular lightning (with no visible shadow since the light was frontal).

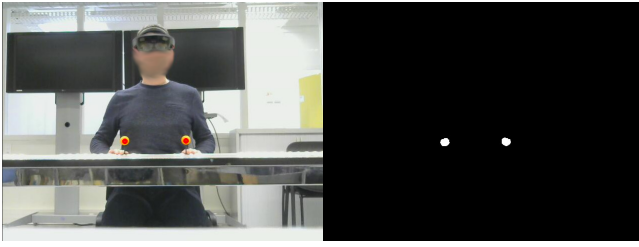


Figure 3: Detection of the spheres' positions with the camera. Left: original image from the camera. Right: Image obtained after applying a color threshold in the LAB color system.

The HMD was calibrated before the experiment for each participant using the HoloLens' built-in calibration procedure. At the beginning of the experiment, the location of the bench was determined using the markers placed on it that were detected using the Vuforia Engine v8.0.10<sup>1</sup>. This position was registered and remained fixed during the whole experiment.

In order to ensure the correct alignment between real and AR content, before each run we asked participants to visually assess that the real and the virtual spheres displayed and placed at the same position were superimposed. If there were any mismatch between the two spheres, the application was reset. This never occurred except when participants took off the headset during the break, which leads us to believe that the calibration was robust enough for our purpose.

Participants were seated on a fixed chair in front of a table. They were asked to sit straight and with their back resting on the back of the chair during the whole experiment. Two spheres, identical to those placed on the bench, were fixed on sliding support on a rail in front of them, see Fig. 1 left. Participants could answer by sliding these *answering spheres*. The distance between the two answering spheres was recorded by a camera placed under the bench, in front of the participant. To compute the distance between both spheres, the camera took a picture and the image obtained was filtered to detect the spheres' centroid positions. To do so, the picture was first converted in the LAB color system to get rid of the luminosity component and then thresholded to only keep the specific orange color of the sphere (see Fig. 3). The whole procedure was implemented using OpenCV 3.2.0. The camera was calibrated using OpenCV's standard calibration method at the beginning of the experiment by placing the spheres at a reference known distance. The resolution of the camera was 1280×720 px which corresponds to a precision of 2 mm on the recorded answer.

### 3.3 Procedure

Participants started by reading and filling out a short form containing written instructions about the experiment. After verbal explanations, they carried out the in-built calibration procedure of the HoloLens. Then, participants performed 10 trials in order to get used to the experimental procedure. Finally, participants had to perform two blocks of 48 trials each (see Sect. 3.4). Participants were allowed to have a break between the two blocks. The procedure for each trial was as follows.

At the beginning of each trial, the moving panel of the occluder was closed and the moving spheres in front of the participant were side by side, regrouped at the center of the rail. The moving panel opened, revealing zero, one or two real spheres, according to the trial condition. Simultaneously, the virtual spheres (if relevant for the trial) were displayed through the HMD. There were two distinct spheres visible at a time. The spheres were visible for five seconds. After that time, the moving panel closed hiding the physical spheres, and, if present, the virtual spheres disappeared.

<sup>1</sup><https://developer.vuforia.com/>

Table 1: Summary of Independent and Dependent Variables.

Independent Variables			
	<i>Observers</i>	29	(random variable)
<b>C1</b>	<i>Distance (cm)</i>	8	25, 30, 35, 40, 45, 50, 55, 60
<b>C2</b>	<i>Type</i>	4	real-real (RR), virtual-virtual (VV) real-virtual (RV), virtual-real (VR)
	<i>Repetition</i>	2	1, 2
Dependent Variables			
<b>D1</b>	<i>Reporting</i>	Distance between the two spheres positioned by the participant	

The participant then had five seconds to move the spheres in front of him to reproduce the same distance as the distance between the spheres he just saw. During this time, a virtual indicator informed the participant of the time left to answer. Once this time was up, the camera recorded the distance between the two spheres. If the participant was still adjusting the distance after five seconds this response was removed from the analysis. We believe this did not have a strong impact on the results as this happened rarely (less than 0.005% of the trials).

The participant had then five seconds to bring the sphere in front of him at their initial position while the spheres of the bench, behind the occluder, was moving to prepare the next trial.

The proposed protocol was designed to ensure that (1) all participants looked at the stimuli the same amount of time, (2) that the stimuli was placed at an optimal viewing distance to minimize the potential accommodation and vergence bias and (3) that the stimuli was visible simultaneously by both eyes without head motions even when the distance between them was the maximum. This design encouraged participants to exhibit a similar exploration behavior (static), thus potentially reducing variability.

In total, the average time per participant, including pre-questionnaires, instructions, calibration, experiment, breaks, post-questionnaire and debriefing, was 45 min. Participants wore the HMD for approximately 30 min.

### 3.4 Experimental Design

We used a mixed-model within-subjects design with the following conditions:

- **C1**: The distance between the two stimuli. Eight possible values were chosen: 25 cm, 30 cm, 35 cm, 40 cm, 45 cm, 50 cm, 55 cm, 60 cm.
- **C2**: The type for each stimulus, which could be both real, both virtual, or one real and one virtual (*RR*, *VV*, *RV*, *VR*). In the *one real/one virtual* condition, in order to minimize potential biases due to the placement of the "real" sphere (placed on the left or on the right), the placement of the real sphere was counterbalanced. Therefore, for the real-real and virtual-virtual conditions, the number of trials was also doubled to ensure the same number of repetitions for all **C2** conditions. We were expecting that the effect of the side was not significant, but we included it in the design for completeness.

The distance between the participant and the stimuli (2.1 m) was chosen to be in the comfort zone of the HoloLens<sup>2</sup>. This specific distance should limit potential biases due to a mismatching focal distance between the displayed virtual objects and the real objects. In addition, since the horizontal augmented FOV of the HoloLens is 30°, **C1** was chosen to ensure that both spheres could be simultaneously on the user's field of view.

In summary, participants were presented with 96 trials: 8 distances (**C1**) × 4 stimuli (**C2**) × 2 (for asymmetrical cases) × 2

<sup>2</sup><https://docs.microsoft.com/en-us/windows/mixed-reality/comfort>

blocks. For each block, the order of **C1** and **C2** were fully randomized to minimize potential order effects. The only dependent variable was the reported exocentric distance **D1**.

Following previous research, and based on our experimental design, our main research hypotheses were the following:

- **H1** Exocentric distance perception is the most accurate for the real/real condition.
- **H2** Exocentric distance perception for the virtual/virtual is underestimated when compared to the real/real condition.
- **H3** Exocentric distance perception for the real/virtual and virtual/real conditions is in-between the both real/real and virtual/virtual conditions.

### 3.5 Statistical Analysis

The reported distance was analyzed using general linear mixed model. The model included the independent variables **C1** and **C2** as fixed effects and the *participant* factor as a random effect. Tukey post-hoc tests ( $\alpha > 0.05$ ) were done when needed using the Bonferroni correction for multiple comparisons. The degrees of freedom were corrected using the Greenhouse Geisser correction method when the sphericity assumption was violated. The statistical analysis was carried out using the R software using the function *aov\_ez* of the *afex* package. The functions *emmeans* (*multcomp* package) and *glht* (*multcomp* package) were used for pairwise comparisons.

The statistical analysis only considered the relative distance estimates, computed as the percentage between the real distance between stimuli and the estimated distance from participants.

### 3.6 Results

The general linear mixed model of **C1** and **C2** versus the relative distance estimation showed a significant effect for the placement of the stimuli, **C2** ( $F(1.99, 51.63) = 8.89; p < 0.001; \eta_p^2 = 0.25$ ), but not for the distance factor **C1** ( $F(2.66, 69.08) = 2.16; p = 0.11; \eta_p^2 = 0.08$ ), neither for the interaction term **C1**  $\times$  **C2** ( $F(10.68, 277.56) = 1.49; p = 0.14; \eta_p^2 = 0.05$ ). In order to further study the effect of **C2**, we conducted post-hoc tests to detect significant pairwise differences. Results showed three significant pairwise differences, see Fig. 4, namely:

- stimuli were perceived as significantly closer in the *VV* condition compared to the *RR* condition ( $p = 0.017$ ),
- stimuli were perceived as significantly closer in the *RV* condition compared to the *VR* condition ( $p = 0.014$ ),
- stimuli were perceived as significantly closer in the *VV* condition compared to the *VR* condition ( $p < 0.001$ ).

The significant pairwise difference *RR* vs. *VV* did not support **H1** but it supported **H2**, the estimates for the *RR* condition were less accurate, but estimates for the *VV* condition were significantly smaller. In addition to the main differences among visual conditions, we observed a consistent and noticeable overestimation estimates for all visual conditions, above 25% in overall (cf. Fig. 4). In addition, the non-significant main effect from **C1** showed that the similar relative estimates were observed no matter the distance between stimuli.

Regarding **H3**, results were inconclusive as they showed an asymmetry for the *RV* and *VR* conditions. Our results show that the stimuli on the *VR* condition were perceived as significantly farther away than stimuli of the *RV* as well as the *VV* condition. We did not expect that the position of the virtual sphere would play such an important role, and initially assumed that both *RV* and *VR* would give rise to comparable results.

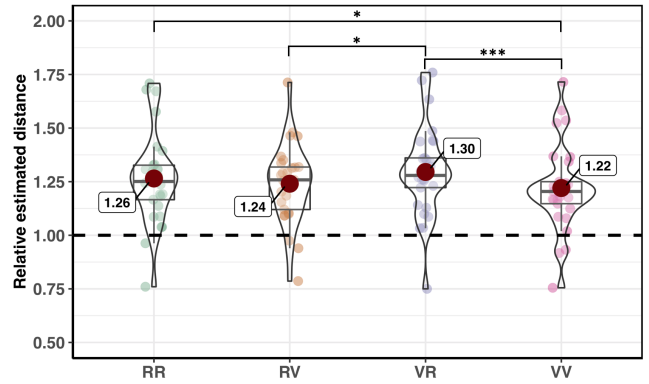


Figure 4: Combined box and violin plot of the relative distance estimation errors averaged for all distances across the four visual conditions: *real-real*, *real-virtual*, *virtual-real* and *virtual-virtual*. The red dot represents the mean value.

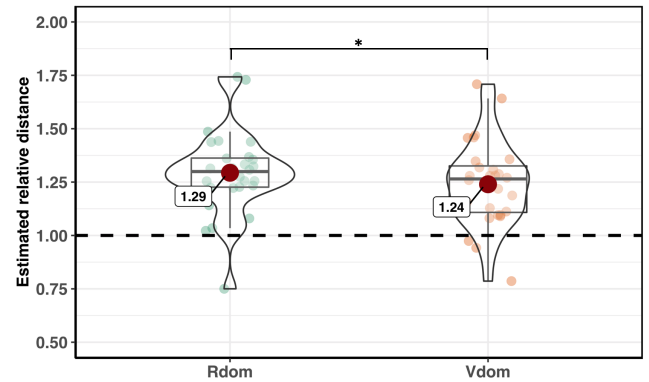


Figure 5: Combined box and violin plot of the relative distance estimation errors for the *Vdom* and *Rdom* trial groups averaged for all distances and conditions.

In order to further study the asymmetry of both conditions with one real and one virtual object, we conducted a statistical analysis integrating the information of participants' eye dominance into account and only considering *RV* and *VR* trials. We thus grouped the trials into two groups (**C3**):

- *Vdom*: trials for which the virtual object is placed on the same side as the participant's dominant eye.
- *Rdom*: trials for which the real object is placed on the same side as the participant's dominant eye.

The general linear mixed model of **C1** and **C3** versus the relative estimated distance. The results only showed a main significant effect for **C3** ( $F(1, 27) = 6.47; p = 0.02; \eta_p^2 = 0.19$ ), there was no effect for **C1** ( $F(3.57, 96.30) = 2.09; p = 0.10; \eta_p^2 = 0.07$ ) nor the interaction term **C1**  $\times$  **C3** ( $F(4.78, 129.09) = 0.43; p = 0.82; \eta_p^2 = 0.02$ ). Pairwise post-hoc tests showed that *Vdom* estimates were significantly smaller than *Rdom* estimates ( $p = 0.017$ ). Fig. 5 shows the summary of the results.

In addition, we further analyzed the data for each participant, comparing the estimate of the average relative distance for each condition. Results showed that 78% of the participants perceived *Vdom* as smaller than *Rdom* and 88% of the participants perceived *VV* as smaller than *RR*.

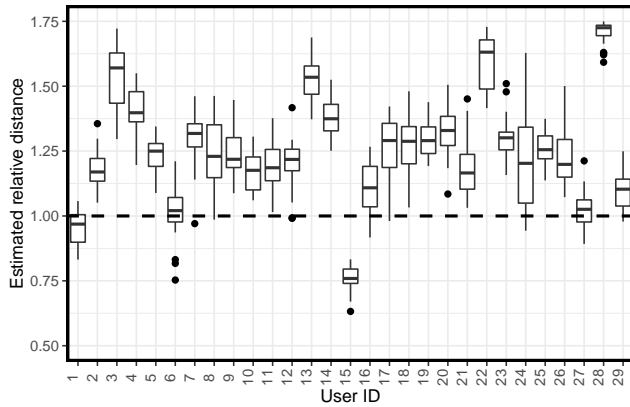


Figure 6: Box plot of the relative distance estimation for each user averaged for all distances and conditions.

Finally, we also explored the potential inter-user variability (see Fig. 6), which showed a relatively strong user variability. For example, the average overestimation for users 3, 13, 22 and 28 was greater than 50% and user 15 had an underestimation of 25% while users 1, 6 and 27 provided estimates close to the real distances.

## 4 DISCUSSION

Results of this study show that the distance between two objects presented in OST AR in the perceptual middle-field are overall overestimated (by around 25.5%) across all four conditions where both objects can be either: *real-real*, *real-virtual*, *virtual-real* and *virtual-virtual*.

When further analyzing those results, we also notice that two virtual objects in OST AR are perceived as significantly closer than two real objects at the same positions. For the farthest distance (over 50 cm between both objects), this effect is equivalent to an underestimation of 4% for the *both virtual* case when compared with the *both real* layout. This effect is in line with previous results of studies showing an underestimation of exocentric distances both in AR [5] and in VR [14], but here we observed it for smaller distances.

One possible explanation of this phenomenon can be linked to the underestimation of egocentric distances. Under the tested condition, if the virtual spheres are perceived closer than the real ones, and since their positions in the visual field is the same, the resulting exocentric distance is perceived smaller.

Even if the VE was carefully designed to provide the same lighting, the limitations of the OST display induced a different representation between the virtual and the real spheres. Therefore, there were clear, noticeable differences in lighting and shading according to the nature (real or virtual) of the sphere [1]. This difference in rendering between real and virtual spheres could thus lead to a bias in perceived distances.

The global overestimation of distances for all the tested conditions is not a usual result. In our case, exocentric distances are reported as 25.5% larger on average. This result is unusual since the majority of studies report a rather accurate exocentric distance perception in the frontoparallel plane, see [12]. However, these results can be linked with the chosen distances between objects as well as the reporting procedure chosen in our experiment. Here, participants had to move two real objects in front of them. Yet, the proprioceptive distance perception is known to be non-homogeneous [8]. Since the reporting task involves hand movement, the answer could be biased by this effect.

Finally, the most surprising result in our study corresponds to the asymmetrical results from the *real-virtual/virtual-real* setups. These results are more difficult to interpret. While we expected for

these two layouts to be symmetrical, this was not the case. We also hypothesized (see **H3**) that the *RV* and *VR* conditions would lead to results in-between those of the *RR* and *VV* conditions. This result is unexpected since previous results, see Lin et al. [18], showed a constant underestimation in the *one virtual/one real* situation.

We thus tried to further analyze this result to understand what could lead to this asymmetry of error estimations based on the position of the virtual stimulus. Since our experimental setup was designed to be completely symmetrical, this effect of the side was not expected and we suggested that it relies on some internal lateralization of the participants. To do so, we grouped trials based on the position of the stimulus in relation with the participants' dominant eye of the subject. Our results showed a significant effect between *Vdom* (virtual object on the side of the dominant eye) and *Rdom* (real object on the side of the dominant eye) with an overestimation in the *Rdom* case of around 5% (see Fig. 5). Therefore, the results are closer to the *real-real* condition when the real object is on the side of the dominant eye, and are closer to the *virtual-virtual* condition when the virtual object is on the side of the dominant eye. As a first explanation of this eye dominance, we supposed that the observer could have focus more one the object on the same side as its dominant eye and consider the other object as symmetrical.

Hand dominance could also be a potential factor explaining this result, yet, only one subject was left-handed, which shows that the effect happened already with a right handed population, ruling out this potential bias.

Another possible explanation for this phenomenon could rely on luminance. Indeed, luminance has been reported as being a strong factor for depth estimation for real objects [1, 4, 7] or more recently in AR [25]. More saturated and brighter objects tend to be perceived as nearer than less saturated and bright objects. Then, a difference of illumination between the screens of the HoloLens could lead to a depth mismatch of the objects and then a perceived exocentric distance biased. Therefore, we measured the respective luminance of the screens by displaying a full screen white rectangle on the HoloLens and taking pictures of the screens using a camera (Point Grey Flea3 FL3-U3-32S2C). We then calculated the average brightness of the central part of the screens (where the sphere was displayed). We found that the right screen was brighter than the left one with a contrast value of 14%. However, since the experiment was designed in way that both spheres were always visible by both eyes at the same time, the potential effect of luminance in that case should also be counterbalanced. Moreover, the HoloLens being an OST HMD, the relative illumination of the augmented objects is background-dependent and can vary because of the head movements of the observer.

As of now, we cannot provide a specific explanation for this phenomenon and further studies are required to better analyze it.

### 4.1 Limitations and Future Work

Our protocol was focused on absolute distance perception. However, our results – and in particular the specific case of the *real-virtual* and *virtual-real* layouts – were difficult to interpret. New experiments, more focused on this specific cases are needed to assess the similarities of distance perception under those conditions. In particular, two-alternative forced choice (2AFC) reporting protocols could be used since they are known to be more accurate when assessing the existence of potential biases.

The present study was conducted with a specific OST HMD: the Microsoft HoloLens. This device has particular characteristics, one of the most important being its limited FOV. This characteristic constrained our experimental design and the position tested for the virtual sphere, since they had to be both displayed. The focal distance of the HMD is also an important factor which should be studied. Then, new experiments can be conducted using other OST HMDs to provide additional views on exocentric distance perception

in AR. The HMD's display technique can also vary, by using a VST HMD, or even with projectors in the case of Spatial Augmented Reality (SAR).

As stated by the literature, space perception varies according to the distance field considered [13]. Here, we focused on one specific distance but exocentric distance perception should be studied for other distance fields. This factor has also to be studied combined with the focal distance of the HMD, as discussed before.

The visual aspect of the considered objects can also influence the results. In this study, we focused on very salient objects, with known size and simple shape. However, the shape and size of the presented objects could vary. Two different objects can also be presented to the observers to evaluate the influence of the relative size of the objects in AR. Moreover, the virtual rendering and shading methods can be changed, in particular to observe the impact of the rendering quality on exocentric distance perception. Lastly, augmented objects – with mixed real and virtual parts – could also be used.

The layout of the objects was here limited to two objects placed in the frontoparallel plane. This layout can be extended by adding other objects to evaluate the impact of a virtual or real object placed between the two original ones. Occluders or objects in other planes can also be used. Finally, the objects' layout can also be modified by rotating the plane where the objects are placed. Indeed, we could not only consider extending our study on exocentric distance perception in the sagittal plane but also on all the possible angles between those two (frontoparallel and sagittal) layouts. This is motivated by previous evidence (see [19]) that physically equal distances in a sagittal plane on the ground are perceptually underestimated relative to those in a frontoparallel plane, even under full-cue conditions. We would expect that placing virtual objects outside of the frontoparallel plane would lead to different exocentric distance estimation.

## 5 CONCLUSION

In this paper, we studied exocentric distance perception in OST AR and compared how distances are perceived between two real objects, two virtual objects and one real and one virtual objects (in two symmetrical layouts). We designed an experimental setup using a Microsoft HoloLens HMD and a self-made bench to precisely place real and virtual objects. Participants had to reproduce exocentric distances between two visible (real and/or virtual) spheres by sliding two similar spheres placed in front of them.

Results of our study show that exocentric distance perception is underestimated for two virtual objects when compared with two real objects. For distances between 50 and 60 centimeters, two virtual objects are perceived as being 4% closer one from another than two real objects. Unexpectedly to us, the layouts with one real and one virtual object give rise to asymmetrical results. This phenomenon was unexpected and, when trying to explain it by correlating the position of the stimuli with the participant's dominant eye, a significant effect was found which suggests that the dominant eye leads the distance perception. We believe further experiments are required to better understand this expected result, illustrating the complexity of distance perception in mixed OST AR setups when participants need to make perceptual judgments based on both real and virtual objects.

Considered together, our results suggest that exocentric distance perception of virtual objects is different from real objects in OST AR. In particular, two virtual objects are perceived closer to one another compared with two real objects. This study lays the groundwork to better understand exocentric distance perception in AR and paves the way to future experiments aiming at further characterizing this phenomenon.

## REFERENCES

[1] M. L. Ashley. Concerning the Significance of Intensity of Light in Visual Estimates of Depth. *Psychological Review*, 5(6):595–615, 1898.

- doi: 10.1037/h0068517
- [2] J. A. Aznar-Casanova, E. H. Matsushima, J. A. Da Silva, and N. P. Ribeiro-Filho. Can exocentric direction be dissociated from its exocentric distance in virtual environments? *Perception & Psychophysics*, 70(3):541–550, 4 2008. doi: 10.3758/PP.70.3.541
- [3] M. A. Cidota, R. M. Clifford, S. G. Lukosch, and M. Billinghurst. Using Visual Effects to Facilitate Depth Perception for Spatial Tasks in Virtual and Augmented Reality. In *Adjunct Proceedings of the 2016 IEEE International Symposium on Mixed and Augmented Reality, ISMAR-Adjunct 2016*, pp. 172–177. IEEE, 9 2017. doi: 10.1109/ISMAR-Adjunct.2016.0070
- [4] J. Coules. Effect of photometric brightness on judgments of distance. *Journal of Experimental Psychology*, 50(1):19–25, 1955. doi: 10.1037/h0044343
- [5] A. Dey, G. Jarvis, C. Sandor, and G. Reitmayr. Tablet versus phone: Depth perception in handheld augmented reality. In *ISMAR 2012 - 11th IEEE International Symposium on Mixed and Augmented Reality 2012, Science and Technology Papers*, pp. 187–196. IEEE, 11 2012. doi: 10.1109/ISMAR.2012.6402556
- [6] A. C. Durand and G. M. Gould. A method of determining ocular dominance. *Journal of the American Medical Association*, 55(5):369–370, 7 1910. doi: 10.1001/jama.1910.04330050007004
- [7] M. Farne. Brightness as an indicator to distance: relative brightness per se or contrast with the background? *Perception*, 6(3):287–293, 1977. doi: 10.1068/p060287
- [8] E. D. Fasse, N. Hogan, B. A. Kay, and F. A. Mussa-Ivaldi. Haptic interaction with virtual objects: Spatial perception and motor control. *Biological Cybernetics*, 82(1):69–83, 2000. doi: 10.1007/PL00007962
- [9] J. M. Foley. Depth, size and distance in stereoscopic vision. *Perception & Psychophysics*, 3(4):265–274, 7 1968. doi: 10.3758/BF03212742
- [10] J. M. Foley, N. P. Ribeiro-Filho, and J. A. Da Silva. Visual perception of extent and the geometry of visual space. *Vision Research*, 44(2):147–156, 2004. doi: 10.1016/j.visres.2003.09.004
- [11] M. Geuss, J. K. Stefanucci, S. H. Creem-Regehr, and W. B. Thompson. Can I Pass?: Using Affordances to Measure Perceived Size in Virtual Environments. In *Proceedings of the 7th Symposium on Applied Perception in Graphics and Visualization - APGV '10*, pp. 61–64, 2010. doi: 10.1145/1836248.1836259
- [12] M. N. Geuss, J. K. Stefanucci, S. H. Creem-Regehr, and W. B. Thompson. Effect of viewing plane on perceived distances in real and virtual environments. *Journal of Experimental Psychology: Human Perception and Performance*, 38(5):1242–1253, 10 2012. doi: 10.1037/a0027524
- [13] I. P. Howard. *Seeing in depth, Vol. 1: Basic mechanisms*. University of Toronto Press, Toronto, ON, Canada, 2002.
- [14] J. W. Kelly, W. Hammel, L. A. Sjolund, and Z. D. Siegel. Frontal extents in virtual environments are not immune to underperception. *Attention, Perception, and Psychophysics*, 77(6):1848–1853, 8 2015. doi: 10.3758/s13414-015-0948-8
- [15] R. V. Kenyon, D. Sandin, R. C. Smith, R. Pawlicki, and T. Defanti. Size-Constancy in the CAVE. *Presence: Teleoperators and Virtual Environments*, 16(2):172–187, 4 2007. doi: 10.1162/pres.16.2.172
- [16] N. Kudoh. Dissociation between Visual Perception of Allocentric Distance and Visually Directed Walking of its Extent. *Perception*, 34(11):1399–1416, 11 2005. doi: 10.1068/p5444
- [17] C. A. Levin and R. N. Haber. Visual angle as a determinant of perceived interobject distance. *Perception & Psychophysics*, 54(2):250–259, 3 1993. doi: 10.3758/BF03211761
- [18] C. J. Lin, B. H. Woldegiorgis, D. Caesaron, and L. Y. Cheng. Distance estimation with mixed real and virtual targets in stereoscopic displays. *Displays*, 36:41–48, 1 2015. doi: 10.1016/j.displa.2014.11.006
- [19] J. M. Loomis, J. A. Da Silva, N. Fujita, and S. S. Fukusima. Visual Space Perception and Visually Directed Action. *Journal of Experimental Psychology: Human Perception and Performance*, 18(4):906–921, 1992. doi: 10.1037/0096-1523.18.4.906
- [20] J. M. Loomis and J. M. J. Knapp. Visual perception of egocentric distance in real and virtual environments. In *Virtual and adaptive environments: Applications, implications, and human performance issues*, number 11, pp. 21–46. CRC Press, 6 2003. doi: 10.1201/9781410608888
- [21] J. M. Loomis, J. W. Philbeck, and P. Zahorik. Dissociation between



- location and shape in visual space. *Journal of experimental psychology. Human perception and performance*, 28(5):1202–12, 10 2002.
- [22] J. F. Norman, J. T. Todd, V. J. Perotti, and J. S. Tittle. The Visual Perception of Three-Dimensional Length. *Journal of Experimental Psychology: Human Perception and Performance*, 22(1):173–186, 1996. doi: 10.1037/0096-1523.22.1.173
- [23] G. Pointon, C. Thompson, S. Creem-Regehr, J. Stefanucci, M. Joshi, R. Paris, and B. Bodenheimer. Judging action capabilities in augmented reality. In *Proceedings of the 15th ACM Symposium on Applied Perception - SAP '18*, pp. 1–8. ACM Press, New York, New York, USA, 2018. doi: 10.1145/3225153.3225168
- [24] J. Predebon. Spatial judgments of exocentric extents in an open-field situation: Familiar versus unfamiliar size. *Perception & Psychophysics*, 50(4):361–366, 7 1991. doi: 10.3758/BF03212228
- [25] G. Singh, S. R. Ellis, and J. E. Swan. The Effect of Focal Distance, Age, and Brightness on Near-Field Augmented Reality Depth Matching. *IEEE Transactions on Visualization and Computer Graphics*, 2018. doi: 10.1109/TVCG.2018.2869729
- [26] J. K. Stefanucci, S. H. Creem-Regehr, W. B. Thompson, D. A. Lessard, and M. N. Geuss. Evaluating the Accuracy of Size Perception on Screen-Based Displays : Displayed Objects Appear Smaller Than Real Objects. *Journal of Experimental Psychology: Applied*, 21(3):215–223, 2015.
- [27] N. Sugano, H. Kato, and K. Tachibana. The effects of shadow representation of virtual objects in augmented reality. *Proceedings - 2nd IEEE and ACM International Symposium on Mixed and Augmented Reality, ISMAR 2003*, pp. 76–83, 2003. doi: 10.1109/ISMAR.2003.1240690
- [28] J. E. Swan, G. Singh, and S. R. Ellis. Matching and Reaching Depth Judgments with Real and Augmented Reality Targets. *IEEE Transactions on Visualization and Computer Graphics*, 21(11):1289–1298, 2015. doi: 10.1109/TVCG.2015.2459895
- [29] J. A. Thomson. Is continuous visual monitoring necessary in visually guided locomotion? *Journal of Experimental Psychology: Human Perception and Performance*, 9(3):427–443, 1983. doi: 10.1037/0096-1523.9.3.427
- [30] J. Vishton and P. Cutting. Perceiving Layout and Knowing Distances: The Integration, Relative Potency, and Contextual Use of Different Information about Depth. *Perception of Space and Motion*, 22(5):69–117, 1995. doi: 10.1016/B978-012240530-3/50005-5
- [31] D. Waller. Factors affecting the perception of interobject distances in virtual environments. *Presence: Teleoperators and Virtual Environments*, 8(6):657–670, 12 1999. doi: 10.1162/105474699566549
- [32] C. Wartenberg and P. Wiborg. Precision of Exocentric Distance Judgments in Desktop and Cube Presentation. *Presence: Teleoperators and Virtual Environments*, 12(2):196–206, 4 2003. doi: 10.1162/105474603321640941
- [33] B. H. Woldegiorgis and C. J. Lin. The accuracy of distance perception in the frontal plane of projection-based stereoscopic environments. *Journal of the Society for Information Display*, 25(12):701–711, 12 2017. doi: 10.1002/jsid.618