Depth Perception in Augmented Reality: The Effects of Display, Shadow, and Position

Haley Adams* Vanderbilt University, USA
Jeanine Stefanucci† University of Utah, USA
Sarah Creem-Regehr‡ University of Utah, USA
Bobby Bodenheimer§ Vanderbilt University, USA

ABSTRACT

Although it is commonly accepted that depth perception in augmented reality (AR) displays is distorted, we have yet to isolate which properties of AR affect people's ability to correctly perceive virtual objects in real spaces. From prior research on depth perception in commercial virtual reality, it is likely that ergonomic properties and graphical limitations impact visual perception in head-mounted displays (HMDs). However, an insufficient amount of research has been conducted in augmented reality HMDs for us to begin isolating pertinent factors in this family of displays. To this end, in the current research, we evaluate absolute measures of distance perception in the Microsoft HoloLens 2, an optical see-through AR display, and the Varjo XR-3, a video see-through AR display. The current work is the first to evaluate either device using absolute distance perception as a measure. For each display, we asked participants to verbally report distance judgments to both grounded and floating targets that were rendered either with or without a cast shadow along the ground. Our findings suggest that currently available video see-through displays may induce more distance underestimation than their optical see-through counterparts. We also find that the vertical position of an object and the presence of a cast shadow influence depth perception.

Keywords: OST AR, VST AR, distance, perception, shadow, depth, surface contact

Index Terms: I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Augmented Reality; J.4 [Computer Applications]: Social and Behavioral Sciences—Psychology

1 INTRODUCTION

Few comparisons of depth perception between stereoscopic video see-through (VST) and optical see-through (OST) head-mounted displays (HMDs) exist [3, 11, 63]. However, direct comparisons of how different augmented reality (AR) displays influence spatial perception can provide important insights into how the technical tradeoffs between AR HMDs influence perception. Direct comparisons are also beneficial for establishing developer guidelines for where and how virtual objects should be rendered in real spaces to enhance depth perception. As such, in this experiment we evaluated distance perception in two augmented reality displays: the Microsoft HoloLens 2 and the Varjo XR-3 (Figure 1). Although both devices provide augmented reality via head-mounted display (HMD), the technology behind them considerably differs.

The Varjo XR-3 functions both as a pure virtual reality device and a video see-through (VST) augmented reality (AR) display. As a result, it shares similar ergonomic and display properties to contemporary virtual reality displays—although the XR-3 boasts exceptional video see-through capabilities. In contrast, the HoloLens 2 is a lightweight display that relies on optical see-through (OST) technology to create augmented reality. Instead of using a display panel to render virtual overlays, the HoloLens 2 projects overlays onto a plastic shield in front of the viewer's eyes. As a result, in the HoloLens viewers have an unaltered view of the real world, but the augmented field of view in OST AR is much smaller. For reference, a more thorough comparison between the two device specifications can be seen in Table 1.

The unique mechanical properties and rendering approaches of these devices have ramifications for how virtual objects are rendered and integrated into real world scenes. One consequence of employing video feed to capture real world images for VST AR is that the physical displacement between the viewer's eyes and the cameras introduces a misalignment that may cause the depth of the scene to be distorted. Similarly, cameras may introduce optical aberrations, like minification or magnification, to the real world image. In contrast, although OST AR displays do not distort real world images, the use of additive light to render virtual overlays causes virtual objects to appear transparent. The darker the color value; the more transparent the overlay. In certain lighting scenarios, like outside on a sunny day, this may cause pictorial depth cues, like cast shadows, to be less salient or even imperceptible.

In addition to AR device, we carefully selected two attributes of virtual targets to better understand how these attributes influence the perception of absolute distances. The first was the vertical position of a virtual object. In order to make gazing at overlays more comfortable, AR applications may present virtual objects as floating, or vertically displaced above the ground (e.g., Google Maps AR). Perhaps because of this, much of the prior research investigating depth perception in AR has used floating virtual targets for assessment [23,28,34,54,72,73,86]. Yet, the decision to evaluate floating targets may have an undesired effect on people's depth judgments.

The distal horizon as well as the height of an object relative to the ground influences depth perception judgments [64, 65]. As such, the position of an object relative to a surface, like whether the object is floating in air or anchored on the ground, alters where we perceive that object to be positioned in space [67, 70]. Because the human visual system treats floating objects as though they are located on the ground plane (in the absence of information specifying otherwise), floating targets are typically perceived as farther away [29, 31]. In augmented reality, as well, the influence of optical contact on distance perception has been demonstrated in the Microsoft HoloLens 1 [81]. Specifically, Salas-Rosales and colleagues demonstrated that floating virtual targets were perceived as on the ground but farther away in AR when no surface contact information, like cast shadows, was present.

The second factor we evaluated was cast shadow. For floating targets, prior research by Ni et al. has shown that the presence of a cast shadow can mitigate the influence of optical contact in virtual environments [68]. As a result, people give more accurate egocentric distance judgments to floating targets when cast shadows are present. Motivated by these findings, a growing body of depth perception research in augmented reality has looked at the effects of cast shadow

---

*e-mail: haley.a.adams@vanderbilt.edu
†e-mail:jeanine.stefanucci@psych.utah.edu
‡e-mail:sarah.creem@psych.utah.edu
§e-mail: bobby.bodenheimer@vanderbilt.edu

on floating targets [23, 28, 34]. However, none of this prior research has looked at the effects of cast shadow for targets placed on the ground. This makes it difficult to extrapolate the results of this prior work to general design guidelines for AR applications, which may be expected to place virtual objects at various heights.

Further, all prior research on cast shadows has relied on perceptual matching, or a *relative* (comparison between two extents) measure, to evaluate the effect of shadows on depth perception. Relative measures are different from *absolute* measures of distance, which rely on stored metrics for estimation rather than comparisons across extents. Thus, absolute measures of distance allow for an understanding of participants' perceived scale of the environment in a given metric (i.e., meters or distance walked) whereas relative measures of depth perception only reveal perception as it compares to other visual extents, so scale is unknown [25, 27]. Evaluations of distance perception can use either measure, but the ability to make estimates with each measure relies on different spatial information [32, 47]. Consequently, it is important to use different measures when studying depth perception since different types of judgments rely on different perceptual representations [19, 75]. We bridge a gap in the literature by evaluating the influence of cast shadows on depth perception using an absolute measure of distance: verbal report.

In summary, we make several contributions in the current study: a) we compare the influence of two commercially available, optical see-through and video see-through AR displays on people's depth perception; b) we isolate the effect of virtual object height on distance perception judgments in AR; c) we extend prior research on the influence of cast shadow by evaluating targets positioned both on and above ground with shadows; and d) we generalize the results of prior research on depth perception in AR by extending the study of shadows in AR to an absolute measure of depth perception.

2 RELATED WORK

2.1 The Importance of Surface Contact

The importance of ground surface for the perception of visual space is perhaps best described by Gibson [29]. In Gibson's seminal work on visual perception, he asserts that the perception of space is impossible without the perception of continuous background. Further, he argues that visual space itself is defined by the layout of surfaces, adjoined surfaces, and entities that are arranged in relation to surfaces [30, 31].

Empirical studies have since shown the importance of surface layout and adjoined surfaces for accurate space perception. Meng and Sedgwick demonstrated that when continuous ground surface is disrupted, the visual system is unable to establish a reliable frame of reference [64, 83]. The visual system, consequently, fails to obtain correct estimates of absolute distance. As such, the accuracy of people's distance perception is disrupted when surface discontinuities along the ground—such as gaps in the floor [84] or changes in texture gradient [64]—are positioned between a viewer and a target.

When cues that link objects to nearby surfaces are absent, individuals judge distance based on *optical contact*—or the location at which the projected image of an object contacts the image of the ground beneath it—to determine position in space [64, 67, 70]. As a result, in the absence of cues specifying that a target is above the ground, distance judgments to targets positioned above the ground are perceived as on the ground but farther away [67, 81]. This phenomenon has been demonstrated in both real [77, 78] and virtual environments [13, 65] within action space, which ranges between 2m and 30m [20]. More recently, Salas-Rosales et al. [81] have confirmed this effect in augmented reality. Specifically, Salas-Rosales and colleagues demonstrated that floating targets are perceived as farther away than grounded ones in an optical see-through augmented reality (OST AR) display, the Microsoft HoloLens 1. However, cues that link objects to the ground plane—like cast shadows and interreflections [92]—can mitigate this effect.

Table 1: Device specifications for the two AR displays are presented[1]. The Varjo XR-3 has differing maximum resolutions for focus and peripheral areas of the display, while the Microsoft HoloLens 2 has one global resolution.

Display Specifications	Varjo XR-3	HoloLens 2
field of view (FoV)	115 × 90°	43 × 29°
peripheral: resolution per eye (pixels)	2880 × 2720	2048 × 1080
peripheral: pixel density (PPD)	30	47
focus: resolution per eye (pixels)	1920 × 1920	2048 × 1080
focus: pixel density (PPD)	70	47
refresh rate (Hz)	90	60
weight (g)	980	566
price (USD)	5500	3500

Figure 1: The Varjo XR-3 (left) is a video see-through (VST) augmented reality display. The HoloLens 2 (right) is an optical see-through augmented reality display.

For depth judgments to targets in space, optical contact may be a correct indicator that an object is in physical contact with the ground, but this is not the case when objects are floating above the ground. It is important to consider how the vertical position of a virtual object can influence people's distance judgments when interpreting prior research on depth perception in AR. While some of this work has been conducted with targets positioned on the ground [27, 42, 43, 45, 76, 88], much more depth perception research has been conducted with floating targets [11, 18, 21, 23, 28, 33, 54, 63, 72, 73, 85, 86, 89].

Few AR studies have considered how the height of a virtual object above the ground (e.g., whether the object is grounded or floating) can influence people's distance judgments. Dey et al. [22] found that height in the visual field influenced people's depth judgments in a mobile AR study. And Kytö et al [54] revealed that people's confidence in depth judgments was worse for floating targets that were higher above the ground (i.e., 1.0m versus 0.5m above the ground) in a VST HMD.

Yet, in the last few years, research published by Salas-Rosales et al. [81] as well as Hertel et al. [34] has provided evidence that depth judgments in AR differ when target objects are presented as floating or along the ground. Specifically, both Salas-Rosales and Hertel provided evidence that people perceive floating objects as farther away than those presented on the ground when cast shadows were not rendered in AR. These findings support prior research in both real and virtual environments that shows people to rely on optical contact information to make depth judgments when surface contact cues are absent [64, 67, 70] (See Section 2.1). The importance of surface contact information for floating objects is further reinforced by recent AR research by Adams et al. [2] in which people were less confident in surface contact judgments to target objects when cast shadows were removed.

---

[1] https://vr-compare.com, https://varjo.com/products/xr-3/, https://www.microsoft.com/en-us/hololens/hardware.

2.2 The Impact of Cast Shadows

Although shadows may seem inconspicuous, they play a crucial role in visual perception by providing us information about size [16, 95, 99] and shape [17, 48]. But perhaps the most important role that shadows play, at least for the current work, is their contribution to our perception of spatial layout. Given the importance of surface layout posited by Gibson [29], it should not be surprising that shadows, which are cast from an object onto a surface, add invaluable information to our understanding of spatial layout. Specifically, shadows provide an indication about where objects are positioned in space by creating points of contact between objects and adjacent surfaces [36, 58, 92].

In augmented reality, rendering cast shadows can be challenging, especially in optical see-through (OST) displays that rely on additive light for rendering. In these displays, the darker the color value, the more transparent a rendered color becomes until it becomes completely transparent when black. Because of this, much of the prior research investigating cast shadows in AR has been directed towards how to best render them in OST AR displays [35, 39, 49, 60, 61] or how new display technology may be constructed to allow for subtractive rendering in optical see-through displays [40, 44].

For both video and optical see-through displays—and, indeed, any graphical device on a computational budget—rendering lighting effects can be expensive. Due to this and due to the dearth of commercially available AR HMDs prior to 2016, few studies in immersive AR have examined the effect of shadows on depth perception [2, 23, 28, 34, 87]. Diaz et al. [23], Gao et al. [28], and Hertel et al. [34] have all found that people's accuracy in relative depth judgments improves when floating targets are rendered with cast shadows.Hertel et al. [34] advanced this research a step further by comparing relative depth judgments for floating targets with cast shadows to grounded targets without shadows. However, because Hertel and colleagues did not include a condition where objects were rendered on the ground with a shadow, as well, it is impossible to interpret the relationship between cast shadow and height above the ground with this study alone.

To understand the results of this prior research, in the current work, we evaluate distance perception to floating and grounded targets—with and without shadows.

2.3 Distance Perception in Augmented Reality HMDs

There are several factors that may play a role in immersive AR, including the depth cues provided, the distances evaluated, and the type of display evaluated. At present, it is difficult to draw reliable connections between distance estimation results for OST and VST AR displays because of differences between devices or experimental protocols. Direct comparisons between devices may provide important insights into how the technical tradeoffs between AR HMDs influence perception. At present, few studies directly compare perception between different AR head-mounted displays [3, 8, 63, 73]. Generalizations from these studies to consumer level devices may be limited because of in-house modifications to the displays in the prior studies, as well. Thus, we believe our current study provides a useful step forward in understanding how the technical differences between AR head-mounted displays influence depth perception.

In OST AR, studies that evaluate depth perception judgments to targets at near distances (distances < 2m) have been variable. While more studies have found that people's depth judgments are overestimated, especially when compared to real world objects [73, 80, 85, 86, 89], a notable amount of work has also found underestimation [62, 63, 72]. However, distance judgments in action space consistently trend toward underestimation [24, 27, 34, 45, 74, 81, 88, 93, 94].

In their absolute distance estimation study, Salas-Rosales et al. [81] found that distance judgment to virtual targets in the Microsoft HoloLens 1 were underestimated by 15% on average when targets were on the ground and by 7% on average when targets were floating. Only one study thus far has been conducted in the Microsoft HoloLens 2. This work, conducted by Hertel & Steinicke [34], evaluated relative depth judgments to farther distances (15m - 75m) via perceptual matching and they found distances were underestimated by approximately 14.3%, on average. Accordingly, in the current work we anticipate that distances will be underestimated in the Microsoft HoloLens 2.

Less distance perception research has been conducted in video see-through (VST) head-mounted displays [24, 26, 74, 93, 94], and much of it has been conducted in recent years. In addition, the majority of these studies have been conducted with either custom built or retrofitted displays, which complicates comparisons across studies. Jamiy et al [24] and Vaziri et al. [93, 94] both created customized video see-through displays by affixing forward-facing cameras to the front of commercial VR devices. Vaziri modified the nVisor ST50 while Jamiy modified an Oculus Rift DK2. Pfeil et al. [74] evaluated distance perception using a ZED Mini camera attached to the front of the HTC Vive.

Jamiy et al. [24] found that absolute, egocentric distances (as measured with verbal reports and blind walking) in VST AR were underestimated by approximately 20% overall, for both measures. Vaziri et al. [94] then compared the effect of unaltered video input and non-photorealistic video input (via a Sobel filter) on distance perception measured via blind walking. They found that distances in both conditions were underestimated when compared with real world distance estimates. Yet they did not find any significant difference in responses between the two VST AR viewing conditions. People underestimated distances by a dramatic 35%, which the authors speculated was due to participants being encumbered by a heavy backpack computer during the study. In their followup work, Vaziri et al. [93], reported less distance underestimation, approximately 10%, when evaluating distance judgments in an open field. With the ZED Mini, Pfeil et al. [74] assessed distance perception to real targets on the ground via blind throwing. They found that in their VST AR condition, people underestimated targets by 7%, on average.

Interestingly, none of the aforementioned distance estimation studies in VST AR evaluated distance judgments to virtual targets in real world spaces. This approach stands in contrast to OST AR depth perception research, which almost exclusively relies on the study of virtual stimuli. Our current study, to the best of the authors' knowledge, is the first to conduct a depth perception study in a contemporary VST HMD with virtual targets, and we are the first researchers to evaluate depth perception in the Varjo XR-3.

3 EXPERIMENT

In the current experiment, we use verbal report as an absolute measure of distance perception in two AR devices: the Microsoft HoloLens 2 and the Varjo XR-3 (Figure 1). We present targets at various distances within action space, and targets are either presented on or above the ground (Figure 2). Each target object is rendered either with or without a cast shadow. We anticipate that judgments of distance will be underestimated in both AR displays, but that there will be more underestimation in VST AR than in OST AR. Based on prior research on the influence of object height above the ground and

Figure 2: Target spheres were presented on and above the ground, and they were rendered with or without a cast shadow.

Table 2: Independent and dependent variables of experiment

Independent Variables			
	observers	24	(random)
H1	distance	3	3, 4.5, 6
H2	shadow	2	yes, no
H3	height	2	0, 0.2m
H4	display	2	vst ar, ost ar
	repetition	3	1,2,3
Dependent Variables			
distance judgments (meters)			

shadow, we anticipated that floating objects without cast shadows would be perceived as farther away than targets on the ground when shadows were absent. However, when shadows were present, we predicted distance judgments to grounded and floating targets to be similar.

In total, we developed four hypotheses for the current experiment:

H1: Targets will be underestimated in both AR devices.

H2: When a shadow is present, people's distance judgments will become more accurate.

H3: There will be an interaction between shadow and target height, such that there will be a difference between floating and grounded objects without shadows, but no difference when shadows are present.

H4: The video see-through display will induce more distance underestimation than the optical see-through display.

3.1 Materials and Apparatus

The experiment was conducted in a 36 x 26 x 9 ft room that provided a 36 ft linear distance forward for placing targets. A university classroom was reserved throughout the duration of the experiment and tables aligned both sides of the participant during the experiment. Images of the room can be seen in Figure 3.

We conducted the experiment in an optical see-through augmented reality HMD and a video see-through augmented reality HMD. For the optical see-through (OST) augmented reality display condition, we employed the Microsoft HoloLens 2. The HoloLens 2 weighs 566g and has a field of view (FoV) of 43° × 29°. Position tracking in the HoloLens 2 is performed by its native inside out spatial tracking method. For the video see-through (VST) augmented reality condition, we used the Varjo XR-3. The XR-3 weighs 980g and has a FoV of 115° × 90°. For position tracking, the tethered display used the SteamVR 2.0 tracking system in conjunction with the Varjo's native depth sensors, which relied on LiDAR and RGB camera fusion. Both systems automatically computed the user's IPD. A more thorough comparison between the two device specifications can be seen in Table 1.

Applications for both devices were developed in Unity version 2020.3.13f1 with the C# programming language. Shaders to render hard shadows were programmed using a variant of the HLSL language that is compatible with the Unity game engine. The cast shadow shader was developed to render shadows with specified color values. Because the HoloLens 2 is unable to render black, a shadow with a grayscale RGB value of 36 was selected. The same shaders were used for both devices.

3.2 Participants

Twenty-four students and staff from Vanderbilt University were invited as volunteers for this experiment in exchange for 10 USD and 45 minutes of their time. The average age was 28.2 ± 8.75 years (Min: 21, Max: 68). Sixteen volunteers were male and nine were female. All participants experienced both the optical see-through AR display and video see-through AR display for a within-subjects experimental design. Our methods were approved by the local institutional review board, and written consent was obtained from volunteers prior to participation. All participants had normal or corrected-to-normal vision.

3.3 Design

To address our hypotheses, we utilized a 2 (display) × 2 (shadow shading) × 2 (target height) × 3 (target distance) within-subjects factorial design. All conditions were presented to every participant. Distance judgments were obtained through verbal report.

The order that the AR displays were experienced was counterbalanced across participants such that half of the volunteers experienced the HoloLens 2 first and half of the volunteers experienced the Varjo XR-3 first. To reduce potential learning effects between display conditions, participants were moved to the opposite side of the room and rotated 180° before beginning the second part of the experiment with the other AR display. A participant standing on opposite sides of the room with each AR display can be seen in Figure 3.

We selected a sphere to be the virtual target (Figure 2). The virtual sphere measured 20 cm in diameter and was rendered with a middle gray RGB color value of 128. Participants viewed the sphere presented at three distances (3m, 4.5m, and 6m). Spheres were either placed on or above the ground at 0.2m. A height of 0.2m was selected to draw comparisons between the current study and that conducted by Salas-Rosales and colleagues [81], which also presented targets at 0.2m above the ground plane.

Prior research has shown that the angle of a virtual light within a scene can influence distance judgments [23, 28]. Therefore, we positioned a virtual, directional light in the scene that rendered cast shadows, when present, immediately beneath the object. This approach to rendering cast shadows is referred to as "drop shadow" in game development and in prior AR research [23]. For each display, this resulted in twelve unique combinations of stimuli.

Except for device, all other factors (i.e., shadow shading, target height, target distance) were pseudo-randomized so that a participant viewed each unique combination once before experiencing the same combination again. All unique combinations were repeated three times, which resulted in a total number of 36 trials per display. Each participant completed a grand total of 72 trials across both displays. With 24 subjects, a total of 1728 trials were collected overall.

3.4 Procedure

Participants were met at the door of the classroom, where they were given a description of the experiment, an informed consent form, a proof of payment form, and monetary compensation for volunteering to participate in the study. The study followed Covid-19 safety protocols set by the university. All participants were wore face masks and equipment was sanitized between sessions.

Before introducing the volunteer to the augmented reality equipment, the experimenter familiarized the participant with units of distance in an adjacent hallway. Depending on the participant's preference, either metric or imperial units of measure were reviewed using a retractable tape measure. Reviewed distances did not exceed 1 meter or 1 yard. After the participant expressed that they were comfortable with the distance units, the experimenter guided them back into the classroom.

The participant was then outfitted with the first head-mounted display, and the protocol was described to them. The were told that "target object would appear at various distances" along the floor relative to the viewer. Each target object appeared for five seconds before disappearing. At which point, the participant called out the estimated distance to the target. After the experimenter recorded the participant's response, the next trial commenced. The beginning of a subsequent trial was denoted by the sound of a beep. The participant was given no feedback on their performance during the experiment.

Figure 3: A participant views the experiment in the OST AR condition (left) and the VST AR condition (right). For each device, the experiment was conducted on opposite sides of the classroom.

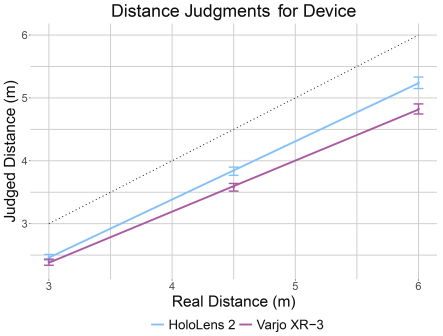

Figure 4: Judged distance vs. real distance for each device condition fit with linear regression. The solid lines represent performance in the HoloLens 2 and Varjo XR-3, while the dotted line represents veridical performance. Error bars are ±1 standard error.

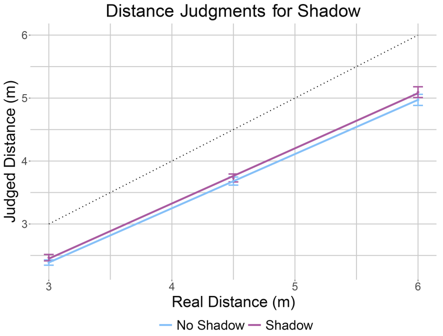

Figure 5: Judged distance vs. real distance for each shadow condition fit with linear regression. The solid lines represent performance for the shadow and no shadow conditions, while the dotted line represents veridical performance. Error bars are ±1 standard error.

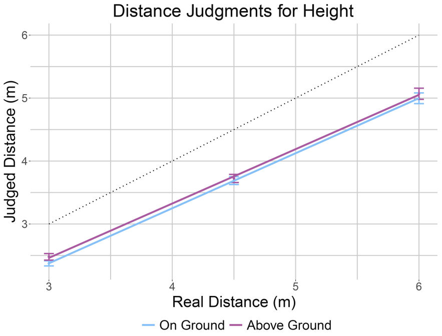

Figure 6: Judged distance vs. real distance for each height condition fit with linear regression. The solid lines represent performance for targets placed on and above the ground, while the dotted line represents veridical performance. Error bars are ±1 standard error.

3.5 Statistical Analysis

Before performing statistical analyses on our data, we converted all recorded distance estimates to meters. The distribution of participants' continuous response data was non-normal, which was exhibited by a positive skew in the data distribution. A Shapiro-Wilk test ($W = 0.943$, $p < 0.001$) and QQ plot inspection of the residuals from our model further revealed that the distribution of the residuals was not normally distributed. Fortunately, linear mixed models are robust to violations of distributional assumptions [82].

Because the population's responses for verbal reports of distance should have a Gaussian distribution, we assume that the underlying distribution of responses from the population is normal for our analyses [5]. Furthermore, to avoid overfitting our predictor values to the current data set, we do not fit our sample data to another distribution [6] nor do we transform our observed data to ensure the results of our analysis are interpretable [14]. Overfitting could compromise our ability to generalize the current results to other samples of the population.

We used a linear mixed-effects model (LMM) to investigate the influence of shadow, target height, and distance on people's distance judgments. Linear mixed models are a form of generalized linear regression that assume a normally distributed dependent variable. They are appropriate for repeated-measures designs because they allow for accounting of both within- and between-participant variability [79]. This is particularly important for examining verbal reports of distance estimates, which can be variable across individuals [52, 57, 66]. LMMs also permit model specification, so our analysis included only the interactions that were hypothesized a priori. This increased our power to detect differences.

Our linear mixed model was programmed using the *lmer* function from the *lme4* library [10] in *R* [38]. We modeled continuous outcomes (verbal distance judgments) for our input variables (predictors). To answer our research questions, device, presence of shadow, distance from viewer, and height of target above the ground were treated as factors. While device (2 levels: HoloLens 2, Varjo XR-3), shadow (2 levels: shadow, none), and height (2 levels: on-ground, above-ground) were treated as categorical factors, target distance was treated as a mean-centered, continuous factor. To answer our experimental hypotheses, the model also included interactions between shadow and target height.

Outside of factors related to our hypotheses, we included experimental block order and the visual context of the room—from participants standing on either the left side of the room or the right side of the room to view targets—in our LMM to better understand how these counterbalanced, experimental factors would influence people's judgments. Both block order (2 levels: first, second) and visual context (2 levels: left, right), were treated as categorical factors. Further, we included interactions between device and experimental block as well as between device and visual context to ensure that these factors did not distort results pertaining to device differences.

To account for individual variability in distance judgment behavior over repeated measures, we included a random intercept (μ_0). We then used Satterthwaite approximation via the *lmerTest* package [53] to calculate significance levels. The general regression equation is depicted in Equation 1 below:

$$\begin{aligned}\Upsilon = {} & B_0 + B_1(\text{device}) + B_2(\text{shadow}) + B_3(\text{height}) \\ & + B_4(\text{distance}) + B_5(\text{order}) + B_6(\text{context}) \\ & + B_7(\text{shadow} \times \text{height}) + B_8(\text{device} \times \text{order}) \\ & + B_9(\text{device} \times \text{context}) + \mu_0\end{aligned} \tag{1}$$

We used the results of this analysis to answer our research questions: 1) whether people's distance judgments would be underestimated (i.e., **H1**); 2) whether peoples distance judgments would improve with the presence of shadows (i.e., **H2** the main effect of shadow); 3) interactions between target height and shadow (i.e., **H3**); and 4) whether distance misperception would be more severe in the Varjo XR-3 than the HoloLens 2 (i.e., **H4** the main effect of device).

3.6 Results

Participants' distance judgments were recorded and statistically analyzed in meters. However, in the following section we also report these values when converted into ratios to facilitate comparisons between the current work and prior research. To create ratios, participants' verbal distance estimates were divided by the actual distances to the target for a given trial. A ratio less than 1 indicates underestimation of distance, and a ratio greater than 1 indicates overestimation. Overall, distance judgments were somewhat variable with a mean distance estimate across participants of 3.722m ($SD = 1.566$, $Min = 0.914$, $Max = 10$). A mean estimate of 3.722m corresponds to a distance ratio of 0.827 or 17.3% underestimation.

3.6.1 The influence of experimental design: Block order and environmental context

Although we counterbalanced our experimental factors, we nonetheless wanted to account for any variance in people's responses that was due to the order of experimental block experienced or due to the visual context provided by the room (i.e., whether the participant viewed the space from the left or right side of the room) by including them within our LMM. Based on prior depth perception studies, effects of experimental block order [27, 100] and visual context may be expected [27, 55, 90, 91]. However, a significant interaction between device and either of these experimental design factors could convolute our planned analysis of device differences.

The analysis revealed a significant effect of block order ($B = 0.126, SE = 0.035, p < 0.001$) and a significant effect of visual context provided by viewing the room at two different locations ($B = 0.353, SE = 0.034, p < 0.001$). The effect of order indicates that people's responses were on average 0.126m farther in the second block of trials. The effect of environmental context indicates that people's responses were 0.353m farther given the environmental context provided by standing at the right side of the room (Figure 3, right) than the context provided by the left side of the room (Figure 3, left). Our analysis did not show an interaction between device and order nor did it show an interaction between device and visual context.

3.6.2 Distance judgments will be underestimated (**H1**)

Participants underestimated distances to all targets by 17.6% on average ($M_{Ratio} = 0.824$, $SD = 0.176$, $Min = 0.305, Max = 2.54$). As shown in Figure 4, participants increased their egocentric distance judgments to virtual targets as the actual distance to the targets increased, supported by the significant main effect of distance ($B = 0.869$, $SE = 0.014$, $p < 0.001$). Participants estimated distances to be approximately 0.87m farther for every meter of increase in actual distance to the sphere, on average.

3.6.3 Distance judgments will be more accurate when a shadow is present (**H2**)

Our analysis supported this prediction. Distance judgments were significantly more accurate when spheres were rendered with a shadow when compared to judgments to spheres without a shadow. This was shown by a significant main effect of shadow ($B = 0.083$, $SE = 0.034$, $p < 0.05$), which indicated that participants estimated distances, on average, were 0.08m farther when a cast shadow was present. Overall, participants underestimated distances to targets rendered with shadows by 16.7% and they underestimated distances to targets without shadows by 18.6%. This relationship is further illustrated by Figure 5.

Table 3: Mean egocentric distance judgments in meters for each device. Values in parentheses are standard errors.

	Each Distance			**All Distances**
Display	3	4.5	6	
All	2.43 (.03)	3.71 (.05)	5.03 (.06)	3.72 (.04)
HoloLens 2	2.47 (.04)	3.83 (.07)	5.24 (.09)	3.85 (.06)
Varjo XR-3	2.39 (.05)	3.78 (.06)	4.82 (.08)	3.60 (.05)

3.6.4 There will be an interaction between shadow and target height (**H3**)

A priori, we anticipated that distance judgments to floating spheres would be similar to those positioned on the ground when shadows were present. Conversely, we predicted that distances to floating spheres would be judged as farther than those positioned on the ground when shadows were absent. Instead, our linear mixed model revealed main effects of both shadows (Section 3.6.3) and height ($B = 0.349$, $SE = 0.174$, $p < 0.05$) with no interaction between the two. When spheres were floating above-ground they were perceived as farther away, with 16.72% underestimation for above-ground spheres and 18.5% underestimation for on-ground spheres. The differences in judgments based on height and on shadows are illustrated by Figures 5 and 6, respectively. Because we found main effects of shadow and height but no interaction between the two, we can conclude that target shadows influenced distance judgments, regardless of target height.

3.6.5 There will be more distance underestimation in the VST AR display than the OST AR display (**H4**)

As shown in Figure 4, egocentric distance judgments to virtual spheres were underestimated in both devices; distances were underestimated by 15.1% in the HoloLens 2 ($M_{Ratio} = 0.849$, $SD = 0.253$, $Min = 0.305, Max = 2.54$), and distances were underestimated by 20.2% in the Varjo XR-3 ($M_{Ratio} = 0.798$, $SD = 0.245$, $Min = 0.333$, $Max = 2.54$). Our statistical analysis indicated that this difference was significant. We found a main effect of device in which participants estimated distances to be 0.25m farther in the HoloLens 2 compared to the Varjo XR-3 ($B = 0.250$, $SE = 0.035\, p < 0.001$). As such, distance underestimation was less severe in the optical see-through display (HoloLens 2) than the video see-through display (Varjo XR-3).

4 DISCUSSION

The perception of scale through AR displays is an important problem that should be understood if AR is to be successfully deployed in applications involving action that takes place over several meters. In this paper, our goal was to understand how AR displays affected the perception of scale, and to understand which characteristics of virtual objects affected that perception when those objects were seen in the context of the real world. We used the Microsoft HoloLens 2, an optical see-through display, and the Varjo XR-3, a video see-through display, to understand how these displays and how characteristics of virtual objects influenced the perception of scale. Both devices are state-of-the-art for their respective display categories, so underlining similar characteristics for both displays may provide us a better understanding of the perceptual issues around AR more broadly.

First, distance judgments across both displays were underestimated, a finding that supports our first hypothesis (**H1**). We found 17.6% underestimation, on average. This result reinforces a growing body of literature evaluating egocentric distance perception in AR head-mounted displays that has found distance estimates in action space (2m - 30m) to be underestimated [24,27,45,74,81,88,93,94]. However, participants were more accurate at estimating distances in the HoloLens 2 than in the Varjo XR-3, with an average of 15.1% underestimation in the HoloLens 2 and 20.2% underestimation in the Varjo XR-3, supporting hypothesis (**H4**).

This latter result confirms prior work in both VR and AR. First, the Varjo XR-3 is heavier than the HoloLens 2 (980 g vs. 566 g), and weight of devices is a known factor in distance underestimation in VR [97]. Second, the field of view of the Varjo XR-3 is narrower than the field of view of the HoloLens 2, which allows nearly unobstructed viewing of the real world scene. Field of view is also a factor in distance underestimation in VR [15,41,42] and AR [74]. Finally, any misalignment of the cameras used in a VST AR display may cause disparities resulting in the depth of the scene being distorted [9]. Likewise, a magnification or minification of the scene seen through these cameras could cause misperception of depth [50]. The latter problems might also exist for graphically displayed objects in an optical see-through display but would not affect the real world objects seen through this type of display.

Additionally, we found a statistically significant effect of shadow presence or absence on distance judgments, confirming (**H2**). But the improvement in distance perception was small, about 2%. This improvement is smaller than prior work in the real world and what our knowledge of graphics would predict [4,23,28,34,37,96]. This finding is important regardless of the size of the effect, because, insofar as we are aware, we are the first to use an absolute measure of distance perception in judging the effect of shadows on distance perception in AR (perceptual matching was used in prior work [23,28,34]). It is important to confirm effects on depth perception through a variety of means and measures to the underlying representations of space.

4.1 Limitations

Our experiment did not find an interaction of height and shadow as we hypothesized (**H3**). The interaction of height and shadow is an effect we hypothesized based on prior work by Salas-Rosales et al. [81] in AR, and this interaction is one predicted by Gibson's ground theory [29]. We did find an effect of height on depth judgments, which is consistent with prior work both in the real world and AR [21,54,69], but the failure of rendered shadows to pin the location of objects down, i.e., the lack of a significant interaction, must be seen as a limitation of our experiment. This limitation should be noted since prior work has demonstrated that what people regard as a shadow in the real world is flexible [59,68].

We conjecture two possible reasons for this lack of an interaction. First, as mentioned previously, the overall effect of shadow on distance underestimation was smaller than we expected *a priori*, particularly based on prior work [81]. It is possible that we did not have enough experimental power to detect an interaction given this. It may be that both the small effect and lack of power are an outcome of using verbal reports. Verbal report measures are common in evaluations of absolute distance perception in both VR and AR [25,27,33,45,52,66] since they can be employed when displays, like the Varjo XR-3, are restrictively tethered and when mobility issues make it impossible for people to perform action based measures [12]. They can also be used to evaluate far distances (e.g., 30m or more) [26,34]. However, verbal report measures can be more variable than other distance measures [7,51,56,98] and they can be susceptible to anchoring effects [71,90,91]. Thus, participants may not make verbal distinctions beyond the nearest 0.25 m (for example) or may repeat common responses.

Second, the choice of a sphere may have made it harder for participants to judge the effect of shadows and ground contact than other shapes that are more commonly used in distance estimation studies in VR and AR, such as a cube [81], traffic cone [46], or hockey puck [15]. Although the use of a sphere with shadows in distance estimation studies has significant antecedents [99], we have prior work that shows prediction of surface contact is harder for a sphere than other shapes (like a cube), and this may have also affected results [1]. Future work should explore these questions more thoroughly.

Another potential limitation of the current work is that we do not directly compare people's distance reports between real and augmented reality targets in the current study. Yet verbal reports often exhibit some degree of underestimation, even in real world studies [7,75]. One way to obtain more accurate comparisons of distance judgments between real and augmented reality targets is to evaluate people's judgments in both to allow for direct comparison [25].

Although our current study provides a foundation for future investigations to evaluate how technical trade-offs between optical see-through and video see-through displays influence depth perception, it is impossible to infer what specific differences in these displays cause dissimilar degrees of underestimation with the current study alone. Part of the reason for this is that there are a large number of hardware and software differences between the two devices. Additional research is required to isolate influencing factors.

4.2 Future Work

Aside from the issues discussed previously, we plan to generalize our findings in terms of both technology (different AR displays) and methods (different ways of evaluating depth perception). In particular, a comparison of the use of verbal reports, blind walking, and perceptual matching to measure distance perception in modern AR displays might address several of the limitations with this paper. Additionally, given the influence of height above the ground found in the the current work, it may be beneficial to conduct a more in-depth investigation of the influence of target height on depth perception in AR. It may be interesting to look at how the unique rendering properties of the two displays influence depth perception, as well, given that OST and VST AR devices integrate virtual and real environment information in unique ways.

4.3 Conclusion

This paper evaluated depth perception in two modern head-mounted AR displays using a verbal report method. We found significant underestimation of distances in both displays, and we found significant differences between the two displays, with the optical see-through display (Microsoft HoloLens 2) having superior performance compared to the video see-through display (Varjo XR-3). We also looked at the effect of shadows and object height on distance estimates. Although we found effects of both shadow and height as perceptual theory and previous work would predict, i.e., that shadows would improve distance estimates and higher objects would be perceived as farther away, the magnitude of the effect caused by virtual objects having shadows was smaller than we expected. We did not find the predicted different effects of object height in the presence versus absence of shadows, which is a fruitful topic for future investigation.

ACKNOWLEDGMENTS

The authors wish to thank Colette Abah and Holly Gagnon for their support with this project. This material is based upon work supported by the Office of Naval Research under grants N00014-18-1-2964, N00014-20-1-2719, and N0014-21-1-2583.

REFERENCES

[1] H. Adams, H. Gagnon, S. Creem-Regehr, J. Stefanucci, and B. Bodenheimer. Stay in touch! shape and shadow influence surface contact in xr displays, 2022. In Prep.

[2] H. Adams, J. Stefanucci, S. H. Creem-Regehr, G. Pointon, W. B. Thompson, and B. Bodenheimer. Shedding light on cast shadows: An investigation of perceived ground contact in ar and vr. *IEEE Transactions on Visualization and Computer Graphics*, pp. 1–1, 2021. doi: 10.1109/TVCG.2021.3097978

[3] J.-g. Ahn, E. Ahn, S. Min, H. Choi, H. Kim, and G. J. Kim. Size perception of augmented objects by different ar displays. In C. Stephanidis, ed., *HCI International 2019 - Posters*, pp. 337–344. Springer International Publishing, Cham, 2019.

[4] B. P. Allen. Shadows as sources of cues for distance of shadow-casting objects. *Perceptual and Motor Skills*, 89(2):571–584, 1999. PMID: 10597594. doi: 10.2466/pms.1999.89.2.571

[5] D. G. Altman and J. M. Bland. Statistics notes: the normal distribution. *Bmj*, 310(6975):298, 1995.

[6] D. G. Altman and J. M. Bland. Statistics notes variables and parameters. *Bmj*, 318(7199):1667, 1999.

[7] J. Andre and S. Rogers. Using verbal and blind-walking distance estimates to investigate the two visual systems hypothesis. *Perception & Psychophysics*, 68(3):353–361, 2006.

[8] G. Ballestin, F. Solari, and M. Chessa. Perception and action in peripersonal space: A comparison between video and optical see-through augmented reality devices. In *2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, pp. 184–189, 2018.

[9] M. S. Banks, J. C. A. Read, R. S. Allison, and S. J. Watt. Stereoscopy and the human visual system. *SMPTE Motion Imaging Journal*, 121(4):24–43, 2012. doi: 10.5594/j18173

[10] D. Bates. lme4: Linear mixed-effects models using s4 classes. r package version 0.999375-34. *http://cran. r-project. org/package=lme4*, 2010.

[11] A. U. Batmaz, M. D. B. Machuca, D. M. Pham, and W. Stuerzlinger. Do head-mounted display stereo deficiencies affect 3d pointing tasks in ar and vr? In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 585–592, 2019.

[12] Z. Bian and G. J. Andersen. Aging and the perception of egocentric distance. *Psychology and aging*, 28(3):813, 2013.

[13] Z. Bian, M. L. Braunstein, and G. J. Andersen. The ground dominance effect in the perception of 3–D layout. *Perception & Psychophysics*, 67(5):801–815, 2005.

[14] J. M. Bland and D. G. Altman. Statistics notes: Transforming data. *Bmj*, 312(7033):770, 1996.

[15] L. E. Buck, M. K. Young, and B. Bodenheimer. A comparison of distance estimation in hmd-based virtual environments with different hmd-based conditions. *ACM Trans. Appl. Percept.*, 15(3):21:1–21:15, July 2018. doi: 10.1145/3196885

[16] R. Casati. Size from shadow: Some informational paths less traveled. *Ecological Psychology*, 28(1):56–63, 2016.

[17] P. Cavanagh and Y. G. Leclerc. Shape from shadows. *Journal of Experimental Psychology: Human Perception and Performance*, 15(1):3, 1989.

[18] M. A. Cidota, R. M. S. Clifford, S. G. Lukosch, and M. Billinghurst. Using visual effects to facilitate depth perception for spatial tasks in virtual and augmented reality. In *2016 IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct)*, pp. 172–177, 2016.

[19] S. H. Creem and D. R. Proffitt. Defining the cortical visual systems:"what","where", and "how". *Acta psychologica*, 107(1-3):43–68, 2001.

[20] J. E. Cutting and P. M. Vishton. *Perception of Space and Motion*, vol. 5 of *Handbook of perception and cognition*, chap. Percieving layout and knowing distances: The integration, relativie potency, and contexual use of different information and depth, pp. 69–117. Academic Press, San Diego, CA, 1995.

[21] A. Dey, A. Cunningham, and C. Sandor. Evaluating depth perception of photorealistic mixed reality visualizations for occluded objects in outdoor environments. In *2010 IEEE Symposium on 3D User Interfaces (3DUI)*, VRST '10, p. 211–218. Association for Computing Machinery, New York, NY, USA, 2010. doi: 10.1145/1889863.1889911

[22] A. Dey, G. Jarvis, C. Sandor, and G. Reitmayr. Tablet versus phone: Depth perception in handheld augmented reality. In *2012 IEEE international symposium on mixed and augmented reality (ISMAR)*, pp. 187–196. IEEE, 2012.

[23] C. Diaz, M. Walker, D. A. Szafir, and D. Szafir. Designing for depth perceptions in augmented reality. In *2017 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 111–122, Oct 2017. doi: 10.1109/ISMAR.2017.28

[24] F. El Jamiy, A. N. R. Chandra, and R. Marsh. Distance accuracy of real environments in virtual reality head-mounted displays. In *2020 IEEE International Conference on Electro Information Technology (EIT)*, pp. 281–287. IEEE, 2020.

[25] I. T. Feldstein, F. M. Kölsch, and R. Konrad. Egocentric distance perception: A comparative study investigating differences between real and virtual environments. *Perception*, 49(9):940–967, 2020.

[26] H. C. Gagnon, L. Buck, T. N. Smith, G. Narasimham, J. Stefanucci, S. H. Creem-Regehr, and B. Bodenheimer. Far distance estimation in mixed reality. In *ACM Symposium on Applied Perception 2020*, pp. 1–8, 2020.

[27] H. C. Gagnon, C. S. Rosales, R. Mileris, J. K. Stefanucci, S. H. Creem-Regehr, and R. E. Bodenheimer. Estimating distances in action space in augmented reality. *ACM Transactions on Applied Perception (TAP)*, 18(2):1–16, 2021.

[28] Y. Gao, E. Peillard, J.-M. Normand, G. Moreau, Y. Liu, and Y. Wang. Influence of virtual objects' shadows and lighting coherence on distance perception in optical see-through augmented reality. *Journal of the Society for Information Display*, 2019.

[29] J. J. Gibson. *The perception of the visual world*. Greenwood Press, Westport, Conn., 1950.

[30] J. J. Gibson. *The implications of experiments on the perception of space and motion*. US Department of Commerce, National Technical Information Service, 1975.

[31] J. J. Gibson. *The ecological approach to visual perception*. Houghton Mifflin, Boston, MA, 1979.

[32] T. Y. Grechkin, T. D. Nguyen, J. M. Plumert, J. F. Cremer, and J. K. Kearney. How does presentation method and measurement protocol affect distance estimation in real and virtual environments? *ACM Trans. Appl. Percept.*, 7:26:1–26:18, July 2010. doi: 10.1145/1823738.1823744

[33] U. Gruenefeld, Y. Brück, and S. Boll. *Behind the Scenes: Comparing X-Ray Visualization Techniques in Head-Mounted Optical See-through Augmented Reality*, p. 179–185. Association for Computing Machinery, New York, NY, USA, 2020.

[34] J. Hertel and F. Steinicke. Augmented reality for maritime navigation assistance-egocentric depth perception in large distance outdoor environments. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*, pp. 122–130. IEEE, 2021.

[35] B. Hu and C. Brown. *Cast Shadows in Augmented Reality Systems*. PhD thesis, University of Rochester, USA, 2005.

[36] H. H. Hu, A. A. Gooch, W. B. Thompson, B. E. Smits, J. J. Rieser, and P. Shirley. Visual cues for imminent object contact in realistic virtual environment. In *IEEE Visualization*, pp. 179–185, 2000.

[37] G. S. Hubona, P. N. Wheeler, G. W. Shirah, and M. Brandt. The relative contributions of stereo, lighting, and background scenes in promoting 3d depth visualization. *ACM Trans. Comput.-Hum. Interact.*, 6(3):214–242, Sept. 1999. doi: 10.1145/329693.329695

[38] R. Ihaka and R. Gentleman. Development core team (2009). r: A language and environment for statistical computing. r foundation for statistical computing, vienna, austria. *URL http://www. R-project. org.(Accessed 16 October 2009)*, 1996.

[39] S. Ikeda, Y. Kimura, S. Manabe, A. Kimura, and F. Shibata. Shadow induction on optical see-through head-mounted displays. *Computers & Graphics*, 91:141 – 152, 2020. doi: 10.1016/j.cag.2020.07.003

[40] Y. Itoh, T. Langlotz, D. Iwai, K. Kiyokawa, and T. Amano. Light attenuation display: Subtractive see-through near-eye display via spatial color filtering. *IEEE Transactions on Visualization and Com-*

puter Graphics, 25(5):1951–1960, 2019. doi: 10.1109/TVCG.2019.2899229

[41] J. A. Jones, E. A. Suma, D. M. Krum, and M. Bolas. Comparability of narrow and wide field-of-view head-mounted displays for medium-field distance judgments. In *Proceedings of the ACM Symposium on Applied Perception*, pp. 119–119, 2012.

[42] J. A. Jones, J. E. Swan, G. Singh, and S. R. Ellis. Peripheral visual information and its effect on distance judgments in virtual and augmented environments. In *Proceedings of the ACM SIGGRAPH Symposium on Applied Perception in Graphics and Visualization*, APGV '11, p. 29–36. Association for Computing Machinery, New York, NY, USA, 2011. doi: 10.1145/2077451.2077457

[43] J. A. Jones, J. E. Swan, G. Singh, E. Kolstad, and S. R. Ellis. The effects of virtual reality, augmented reality, and motion parallax on egocentric depth perception. In *Proceedings of the 5th Symposium on Applied Perception in Graphics and Visualization*, APGV '08, p. 9–14. Association for Computing Machinery, New York, NY, USA, 2008. doi: 10.1145/1394281.1394283

[44] T. Kaminokado, Y. Hiroi, and Y. Itoh. Stainedview: Variable-intensity light-attenuation display with cascaded spatial color filtering for improved color fidelity. *IEEE Transactions on Visualization and Computer Graphics*, 26(12):3576–3586, 2020.

[45] J. Keil, A. Korte, A. Ratmer, D. Edler, and F. Dickmann. Augmented reality (ar) and spatial cognition: effects of holographic grids on distance estimation and location memory in a 3d indoor scenario. *PFG–Journal of Photogrammetry, Remote Sensing and Geoinformation Science*, 88(2):165–172, 2020.

[46] J. W. Kelly, L. A. Cherep, and Z. D. Siegel. Perceived space in the htc vive. *ACM Trans. Appl. Percept.*, 15(1):2:1–2:16, July 2017. doi: 10.1145/3106155

[47] E. Klein, J. E. Swan, G. S. Schmidt, M. A. Livingston, and O. G. Staadt. Measurement protocols for medium-field distance perception in large-screen immersive displays. In *2009 IEEE virtual reality conference*, pp. 107–113. IEEE, 2009.

[48] D. C. Knill, P. Mamassian, and D. Kersten. Geometry of shadows. *JOSA A*, 14(12):3216–3232, 1997.

[49] H. Kolivand, A. H. Zakaria, and M. S. Sunar. Shadow generation in mixed reality: A comprehensive survey. *IETE Technical Review*, 32(1):3–15, 2015. doi: 10.1080/02564602.2014.906860

[50] S. A. Kuhl, W. B. Thompson, and S. H. Creem-Regehr. Hmd calibration and its effects on distance judgments. *ACM Transactions on Applied Perception (TAP)*, 6(3):1–20, 2009.

[51] B. R. Kunz, S. H. Creem-Regehr, and W. B.Thompson. Evidence for motor simulation in imagined locomotion. *Journal of Experimental Psychology: Human Perception and Performance*, 35(5):1458–1471, 2009.

[52] B. R. Kunz, L. Wouters, D. Smith, W. B. Thompson, and S. H. Creem-Regehr. Revisiting the effect of quality of graphics on distance judgments in virtual environments: A comparison of verbal reports and blind walking. *Attention, Perception, & Psychophysics*, 71(6):1284–1293, 2009.

[53] A. Kuznetsova, P. B. Brockhoff, R. H. B. Christensen, et al. Package 'lmertest'. *R package version*, 2(0):734, 2015.

[54] M. Kytö, A. Mäkinen, T. Tossavainen, and P. T. Oittinen. Stereoscopic depth perception in video see-through augmented reality within action space. *Journal of Electronic Imaging*, 23(1):1 – 11, 2014. doi: 10.1117/1.JEI.23.1.011006

[55] J. S. Lappin, A. L. Shelton, and J. J. Rieser. Environmental context influences visually perceived distance. *Perception & psychophysics*, 68(4):571–581, 2006.

[56] J. M. Loomis and J. M. Knapp. *Virtual and Adaptive Environments*, chap. Visual perception of egocentric distance in real and virtual environments, pp. 21–46. Erlbaum, Mahwah, NJ, 2003.

[57] J. M. Loomis and J. W. Philbeck. Measuring spatial perception with spatial updating and action. In R. L. Klatzky, B. MacWihnney, and M. Behrmann, eds., *Embodiment, Ego-Space, and Action*. Taylor & Francis, New York, 2008.

[58] C. Madison, W. Thompson, D. Kersten, P. Shirley, and B. Smits. Use of interreflection and shadow for surface contact. *Perception & Psychophysics*, 63(2):187–194, 2001.

[59] P. Mamassian. Impossible shadows and the shadow correspondence problem. *Perception*, 33(11):1279–1290, 2004.

[60] S. Manabe, S. Ikeda, A. Kimura, and F. Shibata. Casting virtual shadows based on brightness induction for optical see-through displays. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 627–628, 2018.

[61] S. Manabe, S. Ikeda, A. Kimura, and F. Shibata. Shadow inducers: Inconspicuous highlights for casting virtual shadows on ost-hmds. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 1331–1332, March 2019. doi: 10.1109/VR.2019.8798049

[62] J. W. McCandless, S. R. Ellis, and B. D. Adelstein. Localization of a time-delayed, monocular virtual object superimposed on a real environment. *Presence*, 9(1):15–24, 2000.

[63] D. Medeiros, M. Sousa, D. Mendes, A. Raposo, and J. Jorge. Perceiving depth: optical versus video see-through. In *Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology*, pp. 237–240, 2016.

[64] J. Meng and H. Sedgwick. Distance perception across spatial discontinuities. *Perception & Psychophysics*, 64(1):1–14, 2002.

[65] J. C. Meng and H. Sedgwick. Distance perception mediated through nested contact relations among surfaces. *Perception & Psychophysics*, 63(1):1–15, 2001.

[66] P. E. Napieralski, B. M. Altenhoff, J. W. Bertrand, L. O. Long, S. V. Babu, C. C. Pagano, J. Kern, and T. A. Davis. Near-field distance perception in real and virtual environments using both verbal and action responses. *ACM Transactions on Applied Perception (TAP)*, 8(3):1–19, 2011.

[67] R. Ni, M. Braunstein, and G. Andersen. Distance perception from motion parallax and ground contact. *Visual Cognition*, 12(6):1235–1254, Aug. 2005.

[68] R. Ni, M. L. Braunstein, and G. J. Andersen. Perception of scene layout from optical contact, shadows, and motion. *Perception*, 33(11):1305–1318, 2004. PMID: 15693673. doi: 10.1068/p5288

[69] T. L. Ooi, B. Wu, and Z. J. He. Distance determined by the angular declination below the horizon. *Nature*, 414(6860):197–200, 2001.

[70] K. Ozkan and M. L. Braunstein. Background surface and horizon effects in the perception of relative size and distance. *Visual cognition*, 18(2):229–254, 2010.

[71] C. C. Pagano and G. P. Bingham. Comparing measures of monocular distance perception: Verbal and reaching errors are not correlated. *Journal of Experimental Psychology: Human Perception and Performance*, 24(4):1037, 1998.

[72] E. Peillard, F. Argelaguet, J. Normand, A. Lécuyer, and G. Moreau. Studying exocentric distance perception in optical see-through augmented reality. In *2019 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 115–122, 2019. doi: 10.1109/ISMAR.2019.00-13

[73] E. Peillard, Y. Itoh, G. Moreau, J. M. Normand, A. Lécuyer, and F. Argelaguet. Can retinal projection displays improve spatial perception in augmented reality? In *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 80–89, 2020. doi: 10.1109/ISMAR50242.2020.00028

[74] K. Pfeil, S. Masnadi, J. Belga, J.-V. T. Sera-Josef, and J. LaViola. Distance perception with a video see-through head-mounted display. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pp. 1–9, 2021.

[75] J. W. Philbeck and J. M. Loomis. Comparison of two indicators of perceived egocentric distance under full-cue and reduced-cue conditions. *Journal of Experimental Psychology: Human Perception and Performance*, 23(1):72, 1997.

[76] J. Ping, B. H. Thomas, J. Baumeister, J. Guo, D. Weng, and Y. Liu. Effects of shading model and opacity on depth perception in optical see-through augmented reality. *Journal of the Society for Information Display*, 2020.

[77] K. M. Rand, M. R. Tarampi, S. H. Creem-Regehr, and W. Thompson. The influence of ground contact and visible horizon on perception of distance and size under severely degraded vision. *Seeing and Perceiving*, 25(5):425 – 447, 07 Jul. 2012. doi: 10.1163/187847611X620946

[78] K. M. Rand, M. R. Tarampi, S. H. Creem-Regehr, and W. B. Thompson. The importance of a visual horizon for distance judgments under

severely degraded vision. *Perception*, 40(2):143–154, 2011.

[79] S. W. Raudenbush and A. S. Bryk. *Hierarchical linear models: Applications and data analysis methods*, vol. 1. Sage, 2002.

[80] J. P. Rolland, W. Gibson, and D. Ariely. Towards quantifying depth and size perception in virtual environments. *Presence: Teleoperators & Virtual Environments*, 4(1):24–49, 1995.

[81] C. S. Rosales, G. Pointon, H. Adams, J. Stefanucci, S. Creem-Regehr, W. B. Thompson, and B. Bodenheimer. Distance judgments to on- and off-ground objects in augmented reality. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 237–243, 2019.

[82] H. Schielzeth, N. J. Dingemanse, S. Nakagawa, D. F. Westneat, H. Allegue, C. Teplitsky, D. Réale, N. A. Dochtermann, L. Z. Garamszegi, and Y. G. Araya-Ajoy. Robustness of linear mixed-effects models to violations of distributional assumptions. *Methods in Ecology and Evolution*, 11(9):1141–1152, 2020.

[83] H. A. Sedgwick. *The visible horizon: A potential source of visual information for the perception of size and distance*. Cornell University, 1973.

[84] M. J. Sinai, T. L. Ooi, and Z. J. He. Terrain influences the accurate judgement of distance. *Nature*, 395(6701):497–500, 1998.

[85] 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*, 26(2):1385–1398, 2020.

[86] G. Singh, J. E. Swan, J. A. Jones, and S. R. Ellis. Depth judgments by reaching and matching in near-field augmented reality. In *2012 IEEE Virtual Reality Workshops (VRW)*, pp. 165–166. IEEE, 2012.

[87] N. Sugano, H. Kato, and K. Tachibana. The effects of shadow representation of virtual objects in augmented reality. In *Proceedings of the 2nd IEEE/ACM International Symposium on Mixed and Augmented Reality*, ISMAR '03, p. 76. IEEE Computer Society, USA, 2003.

[88] J. E. Swan, A. Jones, E. Kolstad, M. A. Livingston, and H. S. Smallman. Egocentric depth judgments in optical, see-through augmented reality. *IEEE transactions on visualization and computer graphics*, 13(3):429–442, 2007.

[89] 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

[90] M. Teghtsoonian and R. Teghtsoonian. Scaling apparent distance in natural indoor settings. *Psychonomic Science*, 16(6):281–283, 1969.

[91] R. Teghtsoonian and M. Teghtsoonian. Scaling apparent distance in a natural outdoor setting. *Psychonomic Science*, 21(4):215–216, 1970.

[92] W. B. Thompson, P. Shirley, B. Smits, D. J. Kersten, and C. Madison. Visual glue. *University of Utah Technical Report UUCS-98-007*, 1998.

[93] K. Vaziri, M. Bondy, A. Bui, and V. Interrante. Egocentric distance judgments in full-cue video-see-through vr conditions are no better than distance judgments to targets in a void. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*, pp. 1–9. IEEE, 2021.

[94] K. Vaziri, P. Liu, S. Aseeri, and V. Interrante. Impact of visual and experiential realism on distance perception in vr using a custom video see-through system. In *Proceedings of the ACM Symposium on Applied Perception*, SAP '17. Association for Computing Machinery, New York, NY, USA, 2017. doi: 10.1145/3119881.3119892

[95] L. Wanger. The effect of shadow quality on the perception of spatial relationships in computer generated imagery. In *Proceedings of the 1992 symposium on Interactive 3D graphics*, pp. 39–42, 1992.

[96] L. R. Wanger, J. A. Ferwerda, D. P. Greenberg, et al. Perceiving spatial relationships in computer-generated images. *IEEE Computer Graphics and Applications*, 12(3):44–58, 1992.

[97] P. Willemsen, M. B. Colton, S. H. Creem-Regehr, and W. B. Thompson. The effects of head-mounted display mechanical properties and field of view on distance judgments in virtual environments. *ACM Trans. Appl. Percept.*, 6(2), Mar. 2009. doi: 10.1145/1498700.1498702

[98] A. J. Woods, J. W. Philbeck, and J. V. Danoff. The various perceptions of distance: An alternative view of how effort affects distance judgments. *Journal of Experimental Psychology: Human Perception and Performance*, 35(4):1104, 2009.

[99] A. Yonas, L. T. Goldsmith, and J. L. Hallstrom. Development of sensitivity to information provided by cast shadows in pictures. *Perception*, 7(3):333–341, 1978. PMID: 693233. doi: 10.1068/p070333

[100] C. J. Ziemer, J. M. Plumert, J. F. Cremer, and J. K. Kearney. Estimating distance in real and virtual environments: Does order make a difference? *Attention, Perception, & Psychophysics*, 71(5):1095–1106, 2009.