Perceiving Absolute Distance in Augmented Reality Displays with Realistic and Non-realistic Shadows

Bobby Bodenheimer
bobby.bodenheimer@vanderbilt.edu
Vanderbilt University
USA

Haley Adams
haley.a.adams@vanderbilt.edu
Vanderbilt University
USA

Mirinda M. Whitaker
mirinda.whitaker@utah.edu
University of Utah
USA

Jeanine K. Stefanucci
jeanine.stefanucci@psych.utah.edu
University of Utah
USA

Sarah H. Creem-Regehr
sarah.creem@psych.utah.edu
University of Utah
USA

ABSTRACT

Although distance perception to Augmented Reality (AR) objects has been studied for decades, little is known about absolute distance perception with the newest available AR displays. One significant distinction in categories of head-worn AR displays is whether they are optical see-through (OST) or video see-through (VST). These two types of devices have different methods of rendering that could affect the cues available for perceiving distance. Specifically, rendering cast shadows can be challenging, especially in OST displays that rely on additive light for rendering, and there may be alternative shadow shading methods that are equally as effective for conveying cues to depth. The current study tests absolute egocentric distance judgments to targets 3-6 meters away from an observer with two types of shadows, in two types of AR displays, the Hololens 2 (OST) and the Varjo XR-3 (VST). Shadows were realistic cast shadows or non-realistic shadows in the form of a stylized ring placed beneath the object. Participants verbally reported perceived distance to spherical virtual targets presented on or above the ground, viewed through the displays in a real world classroom. We found overall distance underestimation in both devices, but that estimations were more accurate with the Hololens 2 compared to the Varjo XR-3. There was little support for a difference in accuracy of estimations between shadow conditions or position on or above the ground (confirmed by a Bayesian analysis), suggesting that non-realistic shadows may be a good option for providing additional shading cues for depth in AR.

CCS CONCEPTS

• **Human-centered computing** → **Mixed/augmented reality**; **Empirical studies in HCI**; • **Applied computing** → *Psychology*.

KEYWORDS

Distance perception, Augmented reality, Shadows

ACM Reference Format:
Bobby Bodenheimer, Haley Adams, Mirinda M. Whitaker, Jeanine K. Stefanucci, and Sarah H. Creem-Regehr. 2023. Perceiving Absolute Distance in Augmented Reality Displays with Realistic and Non-realistic Shadows. In *ACM Symposium on Applied Perception 2023 (SAP '23), August 05–06, 2023, Los Angeles, CA, USA.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3605495.3605800

1 INTRODUCTION

Augmented reality is increasingly used in important applications that require accurate perception of depth for virtual objects, for example, in simulation [Li et al. 2017] and training [Kaplan et al. 2021]. Modern AR head-mounted displays (HMDs) are divided into two primary types: (1) optical see-through (OST) displays, such as the Hololens and Magic Leap; and (2) video see-through (VST) displays, such as the Varjo XR-3. OST displays work by combining virtual objects and the real world by superimposing the virtual objects on the real world through a semi-transparent mirror or plate of the HMD. Advantages of OST displays are that the real world is always visible, although it may be dimmer due to loss through the semi-transparent mechanism. Disadvantages typically include limited field of view for the virtual display, and virtual objects that are only additive to the scenes in terms of light and contact. VST displays work by taking images from cameras and presenting them inside an HMD display. Advantages of VST displays are that virtual objects and the camera feed of the real world can be more seamlessly combined to create a unified image. Disadvantages include registration of the camera views with the user's perceptual view, and limited resolution.

Applications involving both OST and VST AR require accurate perception of scale and distance. While there has been significant work studying the issue of depth perception in AR [Adams et al. 2022; Diaz et al. 2017; Kuparinen et al. 2013; Pointon et al. 2018a; Rosales et al. 2019; Swan et al. 2006; Vaziri et al. 2021, 2017], there are still unresolved issues. However, how the display technology for virtual objects affects depth perception has not been adequately addressed. Specifically, the rendering systems in OST displays can deprive virtual objects of salient cues for perceiving their depth within a scene. In particular, today's commercially available OST displays face two primary rendering problems. First, the virtual objects they render cannot completely occlude real objects, and, second, the devices cannot attenuate light arriving from the real world [Ikeda et al. 2020]. While some optical see-through devices

---

SAP '23, August 05–06, 2023, Los Angeles, CA, USA

ACM ISBN 979-8-4007-0252-5/23/08…$15.00
https://doi.org/10.1145/3605495.3605800

have been created that remedy such problems [Kiyokawa et al. 2001], they are not commercial devices. The relevant consequence, important for this paper, is that shadows as one normally thinks of them (an obstruction of light), cannot be rendered. However, "shadows" of some perceptual effectiveness can be rendered using the brightness-contrast illusion and rendering patches of gray light [Adams et al. 2021; Diaz et al. 2017; Ikeda et al. 2020]. This paper explores one particular set of methods for rendering cast shadows with an OST device, described in Adams et al. [2021], and compares it to other shading techniques that may offer effective alternatives for denoting distance and ground contact cues. These methods are examined by testing egocentric depth perception to virtual objects in two modern AR displays, an OST device, the Microsoft Hololens 2, and a VST device, the Varjo XR-3. Note that since light can be attenuated on a pixel by pixel basis in a VST device, neither of the limitations described above pertain to this class of AR devices.

In addition to simply varying the device, we varied two attributes of virtual objects to understand how the display characteristics might affect egocentric distance judgments. First, we varied how the shadows were rendered for the virtual objects. Prior work has shown that non-realistically rendered cast shadows can enhance surface contact perception in some cases [Adams et al. 2022, 2021], and ground contact affects depth judgments [Creem-Regehr et al. 2023]. Thus we rendered objects with realistic cast shadows or with non-realistic stylized rings as shadows. Secondly, we placed virtual objects on and above the ground. The height of an object relative to the ground influences how people perceive its distance from them, and shadows are one cue that help determine the height above ground. Thus, when varying shadow types, it seems natural to assess their efficacy by varying the height of objects as well.

In summary, this paper explores the ramifications of different AR display types for different applications by evaluating their differences in egocentric distance perception. We assess perceived distance using a well-known and regarded measure, verbal reports [Andre and Rogers 2006; Loomis and Philbeck 2008]. We evaluate people's ability to judge egocentric distances to virtual targets in *action space*, the space slightly beyond arms reach but within easy walking distance [Cutting and Vishton 1995] — in our case 3 to 6 meters. Our findings are consistent with prior work in showing a difference between the two HMDs. We find no difference in distance perception estimates associated with type of shadow or height off the ground which suggests that any ground cue may be a good option for providing depth information in AR.

2 BACKGROUND

Depth perception in virtual reality has been a topic of significant research for decades, and good reviews exist for this body of work [Creem-Regehr et al. 2023; Kelly 2022]. The problem of depth perception in AR has been studied, but less than in virtual reality [El Jamiy and Marsh 2019]. AR has similarity to virtual reality, but virtual objects are placed in the real world rather than in a completely virtual environment. Ostensibly, many real world cues that convey depth are thus present. However, virtual objects may lack cues and create perceptual conflicts with the surrounding scene, causing ambiguity about their depth. Virtual reality research has consistently shown egocentric distance underestimation, but results in AR are mixed. Some prior work has shown distance underestimation [Adams et al. 2022; Gagnon et al. 2021a; Kytö et al. 2014; Rosales et al. 2019; Swan et al. 2007; Vaziri et al. 2017], but other studies have found accurate or some overestimation [Jones et al. 2008; Pointon et al. 2018b; Swan et al. 2006]. Our assessment is that most recent research using modern devices supports the finding that distance underestimation is occurring in AR when presented through HMDs. This underestimation seems to occur in both VST [Adams et al. 2022; Kytö et al. 2014; Vaziri et al. 2017] and OST [Adams et al. 2022; Jones et al. 2008; Pointon et al. 2018b; Rosales et al. 2019; Swan et al. 2007] devices.

This paper focuses on two particular depth cues, ground contact and shadows. Both are related. Gibson's ground theory of spatial perception emphasizes the importance of shadows in determining whether objects are in contact with the background surface or not [Gibson 1950]. The exact representation of the shadow required by the visual system for this perception to occur is still an open problem [Santos et al. 2018]. If objects are positioned above the ground, they are typically perceived as more distant than objects at the same distance but positioned in contact with the ground [Ni et al. 2005; Rand et al. 2011, 2012]. Using an OST display, Salas-Rosales and colleagues showed that virtual objects positioned above the ground were perceived as farther away that those on the ground [Rosales et al. 2019]. Since shadows typically provide a compelling cue as to whether an object is in contact with the ground surface or above it [Madison et al. 2001; Mamassian et al. 1998], the manipulation of shadow type and height should interact on distance perception. However, Adams and colleagues [2022] manipulated both the presence or absence of shadows together with the height of virtual objects in both OST and VST AR displays, and did not find such an interaction.

Research has also demonstrated that distance perception in AR is improved when virtual objects are rendered with shadows [Diaz et al. 2017; Gao et al. 2019]. The way in which shadows are rendered has been shown to affect depth judgments, with more transparent shadows serving as less reliable depth cues [Adams et al. 2021; Diaz et al. 2017]. Similarly, Gao and colleagues showed that different levels of lighting misalignment between real and virtual lights influence distance judgments, with poorer judgments associated with increased misalignment.

The current work is most similar to that of Adams et al. [2022], and extends the findings of that work. Adams and colleagues examined egocentric distance judgments in the same OST and VST displays as ours, finding distance underestimation. They examined whether the presence of realistic shadows cast by virtual objects in the display affected distance perception, and, like us, they manipulated the height of the virtual object. They found that the presence of shadows affected distance judgments, with distance judgments being more accurate when shadows were present. However, the improvement was smaller than what would be predicted by work done in the real world and virtual reality. The main novelty of the present work is to examine the effect of a non-realistic cue on distance judgments. As mentioned previously, the motivation for the non-realistic cue comes from prior work by Adams and colleagues that suggests such cues can enhance perception of virtual object contact in AR [Adams et al. 2021].

3 EXPERIMENT

The current experiment used verbal reports to test absolute distance perception to virtual targets presented with OST and VST AR devices, varying the type of shadow, distance, and height of the target. We aimed to assess whether a realistic cast shadow and a stylized (ring) shadow would lead to similar distance estimations in the two devices, as well as whether there would be overall differences in distance estimations between device types. We presented the AR targets both on the ground surface and 20 cm above the ground surface in order to generalize shadow effects to circumstances (a floating object) where shadows could serve to enhance perception of ground contact. Based on prior work using the two AR devices for a similar distance perception task [Adams et al. 2022], we predicted that underestimation would result with both devices, but that there would be less underestimation (judgments more similar to the intended distance) with the Hololens 2 compared to the Varjo XR-3. Given prior work that non-photorealistic shadows benefited judgments of surface contact for objects close to the observer [Adams et al. 2021], we expected that our stylized shadow would work at least as equally as well as the realistic shadow in providing a cue to surface location, predicting little effect of shadow on distance estimations. However, we predicted that there may be a difference in the effect of shadow type depending on the device type because of different rendering methods inherent to the device. Finally, although prior work has shown that observers judge distances to be greater to targets hovering above the ground compared to on the ground [Rosales et al. 2019], the addition of effective shadows should provide information for the location of the target above the ground, reducing this effect. Therefore, if both types of shadows are equally effective, then we would expect no difference in judgments due to the height manipulation. We made the following predictions:

- **H1** Distances will be underestimated, but there will be less underestimation (greater accuracy) with the Hololens 2 compared to the Varjo XR-3.
- **H2** Distance estimations will increase as actual target distance increases.
- **H3** (a) Distance estimations will not differ between realistic versus stylized shadows, but (b) shadow type will have a greater influence in the OST versus VST device (shadow x device interaction).
- **H4** Distance estimations will not change with target height.

3.1 Participants

We collected data from 24 participants at [anonymized] university in exchange for 10 USD and 45 minutes of their time. Fifteen participants were female and 9 were male. Mean age of participants was 27 years (range 21–58). Our experimental methods were approved by the local institutional review board, and written consent was obtained from all volunteers prior to participation. All participants had normal or corrected to normal vision

3.2 Materials and Apparatus

We conducted the experiment in a 11 x 7.9 x 2.7 m classroom (reserved for the experiment) that provided an 11 m linear distance forward for placing targets. See Figure 1 for images of the room.

Figure 1: A participant estimating distances to virtual targets in both augmented reality HMDs. Visible on the floor are outlet covers; the virtual targets were not close to these.

We used two augmented reality HMDs: the Microsoft Hololens 2 and Varjo XR-3. The Hololens 2 is an optical see-through (OST) device that weighs 566g and has a field of view (FOV) of 43° × 29°. The graphical display of the Hololens 2 has a resolution of 2048 × 1080 with a resolution of 47 pixels per degree. Position tracking in the HoloLens 2 is performed by its native inside out spatial tracking method. The Varjo XR-3 is a video see-through device (VST) that weighs 980g and has a FoV of 115° × 90°. The XR-3 uses a bionic display with higher resolution in the central part of the display; this region spans 1920 × 1920 with a resolution of 70 pixels per degree. 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.

We used Unity version 2020.3.13f1 with the C# programming language to develop the applications for both devices. 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.

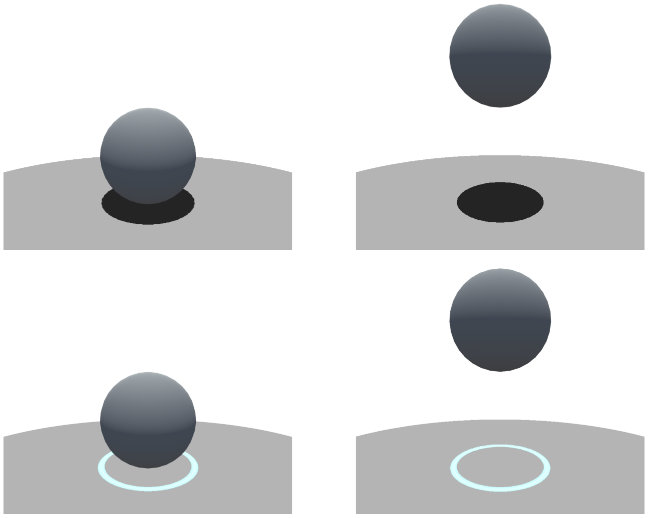

Figure 2: Participants viewed target spheres that were positioned on and above the ground. Targets were rendered with a cast shadow or with a glowing ring on the ground beneath them. Gray surface provided for illustration; the actual ground was shown in the AR devices.

3.3 Design

We used a 2 (device) × 2 (shadow shading) × 2 (target height) × 3 (target distance) within-subjects factorial design such that all conditions were presented to every participant. Half of the participants used the Hololens 2 first and half used with Varjo XR-3 first. We also varied the location in the room that the participants stood to make their judgments, such that they moved to the other side of the room before starting the trials with the second AR device, to reduce possible memory effects that could result with verbal reports.

The shadow shading conditions refer to either a realistic cast shadow or a stylized graphical element along the ground (see Figure 2). For the realistic shadow, we positioned a virtual, directional light in the scene that rendered a cast shadow beneath the object. This method of creating a "drop shadow" has been used in prior AR research on depth perception [Adams et al. 2022; Diaz et al. 2017]. The stylized graphical element was rendered as a glowing white ring underneath the target object. We selected a glowing ring to replace the regular cast shadow, because this kind of graphical element is commonly used in AR applications as well as video games to indicate the position of objects of interest. The glowing ring also provides a more useful analogue to AR user interface elements used in practice.

The target was a virtual sphere that was 20 cm in diameter and rendered with a reflectance of 0.5 (middle gray) (see Figure 2). We presented the sphere at three distances (3m, 4.5m, and 6m), placed on the ground or above the ground at 0.2m. We selected the 0.2m height off the ground based on prior research studying the effects of perceived ground contact in AR [Adams et al. 2022; Rosales et al. 2019]. In particular, Rosales et al. [2019] found an effect of object height on distance perception at these distances without shadows using a cube of similar dimension.

The shadow shading, target height, target distance factors 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 device.

3.4 Procedure

An experimenter met each participant at the door of the classroom, and gave them 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 beginning the AR experiment, the experimenter familiarized the participant with units of distance in an adjacent hallway. Depending on the participant's preference, the experimenter reviewed either metric or imperial units of measure 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 then donned the first HMD, and listened the protocol described to them by the experimenter. The experimenter told them that the "target object would appear at various distances" along the floor relative to the viewer and that they should report the distance that they perceived. Each target object appeared for five seconds before disappearing and then 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 received no feedback on their performance during the experiment.

4 RESULTS

4.1 Data Analysis

Data were analyzed with mixed models, which are appropriate for the nested structure of the data in this experiment. Mixed models allow for the partitioning of variance both within and between participants. All analyses were performed in R [R Core Team 2022]. Mixed models were run using the `lmer` function from the `lme4` package [Bates et al. 2015], and the intraclass correlations (ICC) were calculated with the performance package [Lüdecke 2018].

Due to some of our hypotheses aligning with a null effect, we conducted our analyses in both a frequentist and Bayesian framework. Bayesian models were run using the `rstan`, and `brms` packages [Bürkner 2017; STAN 2018]. For the Bayesian analyses, we report betas, credible intervals, and Bayes factors.[1] A credible interval is a probability statement that the true parameter would lie within an interval a certain percent (e.g., 90, 95, 99) of the time, given the observed data. For example, a 95% credible interval provides a probability statement that given the observed data, the true parameter would fall within the given interval 95% of the time. Bayes factors are mathematically defined by dividing the likelihood of the data

---

[1] Bayes factors were computed using bridge sampling and thus the reported values are approximations as a direct calculation is not possible within the context of multi-level Bayesian models.

under one model/hypothesis by the likelihood of the data under another model/hypothesis, such that a Bayes factor of exactly 1 suggests the data are equally likely under either model/hypothesis.

Results from both frameworks are presented in Table 1 [2]. One note for interpretation, we used a null/uninformative prior in our Bayesian analysis, thus the Bayes factors presented should be interpreted with caution because Bayes factors are highly sensitive to prior specification [Aitkin 1991; Gelman et al. 2013; Grünwald 2000; Liu and Aitkin 2008] [3]. Thus our interpretation in the Bayesian analysis leans more on the information provided by the credible interval, as the posterior distribution summarized by the credible interval is much more stable and less sensitive to the prior, especially as sample size increases.

The primary dependent variable was distance estimates (converted to meters). The intraclass correlation (ICC) for distance estimates was .44, indicating that 44% of the variance in distance estimates was between participants and 56% was within participants. Regardless of condition, on average participants in this experiment underestimated distance by ~27%. All variables were dummy coded in our analyses, which means the intercept represents the average distance estimate when all other variables are set to 0 (i.e., device=hololens, shadow=realistic shadow, height=on ground, distance=3m, order=hololens first, room=hololens on left side of room) and thus all betas are relative to this intercept value.

4.2 Hypothesis 1: Effect of device

On average, participants underestimated distances by ~24% in the Hololens 2 and ~29% in the Varjo XR-3. Hypothesis 1 was supported (via the frequentist analysis) in that distance estimates in the Varjo XR-3 were shorter ($B = -0.44$, SE=.10, $p < .001$), compared to distance estimates in the Hololens 2 (see Figure 3). This means that across participants, distance estimates in the Varjo XR-3 were 44 cm shorter. The Bayesian analysis further confirmed Hypothesis 1 with an identical beta/slope value of -.44 and a 95% credible interval that ranged from −.64 to −.23. While credible intervals are not designed for hypothesis testing [Berger 2006], they do provide intuitive and interpretable estimates of uncertainty. Thus a 95% credible interval ranging from −.64 to −.23 suggests that the range of plausible values for the effect of device span from 23 cm to 64 cm of underestimation in the Varjo XR-3 compared to the Hololens 2. The Bayes factor when comparing the full model to a model with device and all device interaction terms removed was >100 (2526.26), suggesting the data is more than 100 times more likely under the alternative hypothesis than the null hypothesis. However, the effect of device is qualified by a significant device x order interaction (B = .24, SE=.10, p = .02). This suggests there was a larger effect of device ($B = -.44$) when participants ran through the experiment in the Hololens 2 first and the Varjo XR-3 second and smaller effect of device ($B = -.20$) when participants ran through the experiment in the Varjo XR-3 first and the Hololens 2 second.

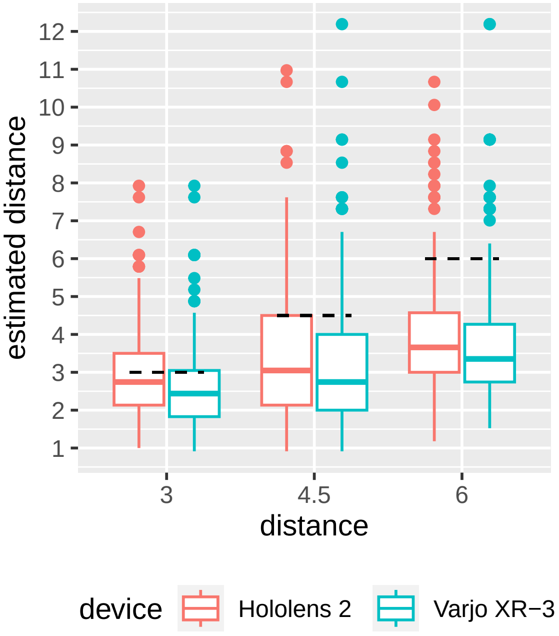

Figure 3: Distance estimates plotted against target distance by device. Middle colored lines are the medians and the box covers the 25^{th} to 75^{th} quartile range. Dashed black lines indicate perfect accuracy/performance.

4.3 Hypothesis 2: Effect of distance

Hypothesis 2 was supported by both the frequentist and Bayesian analyses. In the frequentist analysis, both the 4.5 ($B = 0.58$, SE=0.06, $p < 0.001$) and 6 meter ($B = 1.14$, SE=0.06, $p < .001$) distances were estimated as farther than the 3 meter distance. , as seen in Figures 3-5. This means that on average, at 3 meters (represented by the intercept in our models) participants estimated the sphere to be 2.24 meters away. At 4.5 meters estimates increased by .58 meters, meaning that on average at 4.5 meters participants estimated the sphere to be 2.82 meters away. While at 6 meters estimates increased by 1.14 meters relative to the intercept (3 meters), meaning that on average at 6 meters, participants estimated the sphere to be 3.37 meters away. The Bayesian analysis further confirmed Hypothesis 2 with identical beta/slope values of .58 and 1.14 and 95% credible intervals that ranged from .46 to .70 and 1.01 to 1.26.. The range of plausible values range from 46 cm to 70 cm and 101 cm to 126 cm meaning there is almost certainly a positive relationship between target distance and estimated distance. The Bayes factor for effect of device was >100 (1.27 x 10^{29}), suggesting the data is more than 100 times more likely under the alternative hypothesis than the null hypothesis.

---

[2] During posthoc assumption checking we found that there was some slight heteroscedasticity when plotting residuals vs. fitted values. Thus, we re-ran the analysis using robust standard errors (using the `robustlmm` package in R), while there were some slight changes in standard errors and p-values most changes were in the second or third decimal places and there were no changes in the direction or significance of effects that would change our interpretation. For brevity, only the original analyses are reported

[3] Different priors could lead to a dramatic shift in the Bayes factor (see Liu and Aitkin [2008], for an example).

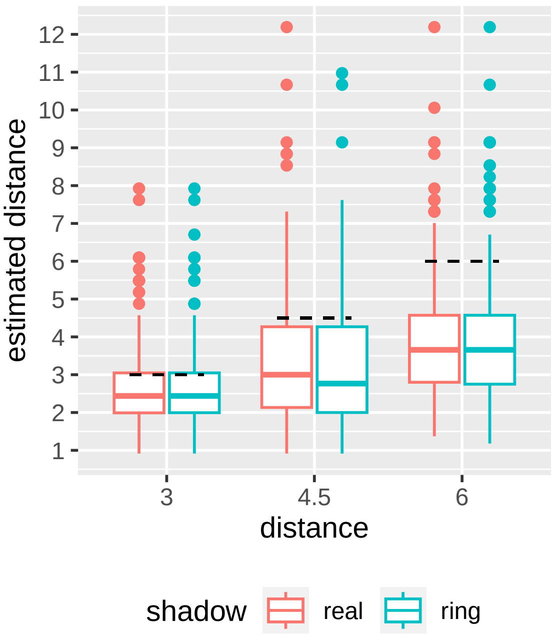

Figure 4: Distance estimates plotted against target distance by shadow type. Middle colored lines are the medians and the box covers the 25^{th} to 75^{th} quartile range. Dashed black lines indicate perfect accuracy/performance.

4.4 Hypothesis 3: Effect of shadows across devices

Hypothesis 3 was partially supported by our frequentist and Bayesian analyses. In the frequentist analysis, our prediction was that there was no effect of shadow (H3(a)) (B = −.01, SE=.09, p = .92), but our prediction failed since the shadow x device interaction term was non-significant (B = .05, SE=.10, p = .52) (H3(b)). The Bayesian analysis resulted in an identical beta/slope value of −.01 and .05 and 95% credible intervals that ranged from −0.18 to 0.16 and −.15 to .25, respectively. Given 0 is near the center of both 95% credible intervals above, we interpret an effect of 0 as plausible for the effect of shadow and the shadow x device interaction. The Bayes factor when comparing the full model to a model shadow and all shadow interaction terms removed was ∼0.01[4], suggesting the data is approximately 99 times more likely under the null hypothesis than the alternative hypothesis. The Bayes factor for the shadow x device interaction effect was ∼0.13, suggesting the data is approximately 8 times more likely under the null hypothesis compared to the alternative hypothesis.

4.5 Hypothesis 4: Effect of height

Hypothesis 4 was supported by both the frequentist and Bayesian analyses. In the frequentist analysis, there was not a significant effect of height (B = .01, SE=.03, p = .88). The Bayesian analysis resulted in an identical beta/slope value of .01 and a 95% credible interval that ranged from −.13 to .15. Given 0 is near the center of the 95% credible interval and the interval is fairly symmetrical, we interpret an effect of 0 as highly plausible. Even if the effect is non-zero it is likely small, as the tails of the 95% credible interval lie around +/- .15. Even at the extremes of plausible values (−.13, .15) this would translate to spheres positioned .2 meters off the ground being either under or overestimated by 13-15 centimeters, with an effect close to 0(.01) being most likely. The Bayes factor for the effect of height was ∼.02, suggesting the data is approximately 62 times more likely under the null hypothesis compared to the alternative hypothesis.

Table 1

	Frequentist	Bayesian
Predictor	Estimate [95% Confidence/Credible Interval]	
Intercept	2.24^{***}[1.56, 2.91]	2.24[1.53, 2.96]
Device	-0.44^{***}[−0.64, −0.24]	−0.44[−0.64, −0.23]
Shadow	−0.01[−0.18, 0.16]	−0.01[−0.18, 0.17]
Height	0.01[−0.13, 0.15]	0.01[−0.13, 0.15]
Distance (4.5 m)	0.58^{***}[0.46, 0.70]	.58[0.46, 0.70]
Distance (6 m)	1.14^{***}[1.01, 1.26]	1.13[1.01, 1.26]
Order	0.75^{*}[−0.07, 1.57]	0.76[−0.07, 1.57]
Room	0.44[−0.33, 1.21]	0.44[−0.39, 1.24]
ShadowXDevice	0.05[−0.15, 0.25]	0.05[−0.15, 0.25]
ShadowXHeight	−0.05[−0.25, 0.15]	−0.05[−0.25, 0.15]
DeviceXRoom	0.10[−0.10, 0.31]	0.10[−0.09, 0.30]
DeviceXOrder	0.24^{**}[0.03, 0.44]	0.24[0.04, 0.44]

Note: *p<0.1; **p<0.05; ***p<0.01

5 DISCUSSION

The current experiment tested whether absolute egocentric distance judgments would be influenced by the type of shadow and the type of AR device used. Previous work had suggested that there would be overall underestimation of distance to AR targets, but how this underestimation is influenced by different shadow rendering techniques in mixed reality devices that use different methods for presenting AR graphics (OST vs. VST) was unknown. In a completely within-subject design, observers verbally reported perceived distance to spherical targets 3-6 meters away, on or above the ground, with realistic and non-realistic shadows, using two different AR devices. Our results mostly supported our hypotheses. **H1:** We found underestimation of distance in both devices but less underestimation with the Hololens 2 (24%) than with the Varjo XR-3 (29%). **H2:** We found the expected effect of distance–as target distance increased, verbal estimates increased. **H3:** We found no difference due to shadow type, and this did not interact with the device type, suggesting that the stylized shadow was as effective as the realistically rendered shadow. **H4:** We found that estimations did not change with location on or above the ground, also providing support for the use of both shadows to specify location relative

---

[4] The "∼" indicates the potential slight variability of this estimate given bridge sampling was used to calculate Bayes factor. A direct calculation is not possible within the context of multi-level models.

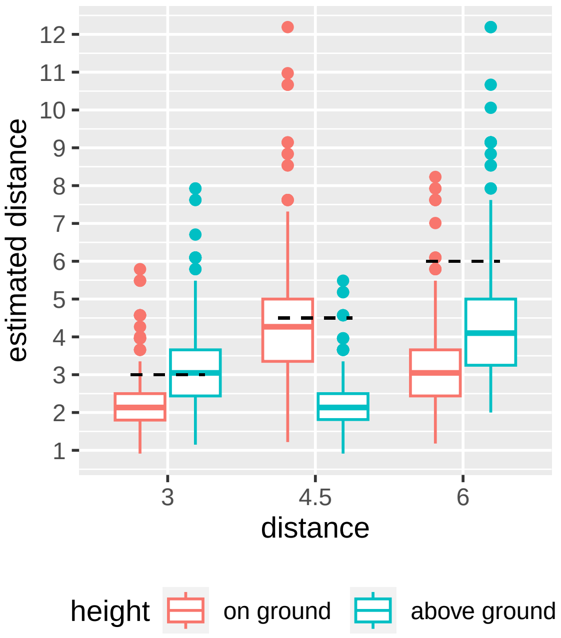

Figure 5: Distance estimates plotted against target distance by height. Middle colored lines are the medians and the box covers the 25^{th} to 75^{th} quartile range. Dashed black lines indicate perfect accuracy/performance.

to the ground plane. We discuss each of these effects (or lack of) further in the context of prior work and potential implications.

Distances were underestimated, as expected, in AR, but participants were more accurate with the Hololens 2 (OST) compared to the Varjo XR-3 (VST) displays. This finding replicated Adams et al. [2022], and is an important result given there are few other studies on distance perception using the relatively new Varjo displays. There are several factors that could possibly contribute the differences found between the devices. One notable factor is the camera-based systems used in VST displays. In the Varjo XR-3, misalignment of the cameras with the eyes could distort depth in the scene [Cattari et al. 2019; Held and Banks 2008; Takagi et al. 2000; Woods et al. 1993]. Another factor could be the different FOVs; a smaller FOV is associated with greater distance underestimation in VR [Buck et al. 2018; Jones et al. 2013; Masnadi et al. 2022, 2021] and AR [Gagnon et al. 2021b; Pfeil et al. 2021]. Although the Hololens 2 has a smaller FOV for the presentation of virtual objects, the viewer sees their full real-world FOV through the OST display, which is larger than the FOV seen through the Varjo XR-3. The weight of the Varjo XR-3 is also greater than that of the Hololens 2, a factor shown to influence distance underestimation (see Kelly [2022] for a recent review).

We found no effect of shadow type on distance estimations, across the two devices. Unlike in Adams et al. [2022], we included shadows on all trials, but varied whether they appeared realistic or stylized as shown in Figure 2. We predicted that the stylized shadow would serve to provide information for ground contact equally well as the realistic shadow, supporting the findings of Adams et al. [2021], who used a different method to assess perceived ground contact at close distances. We also considered the possibility that the shadows would have different effects given the two different display types (i.e., particularly the OST might benefit more from a lighter colored shadow), but we found no evidence for an overall difference or an interaction with device. Why might the stylized, non-realistic shadow be so effective? As in Adams et al. [2021], it could be that the bright ring provided such a salient cue for ground contact that its benefit outweighed any costs of appearing unnatural. The ring used in the current study differs from the solid white shadow used in Adams et al. [2021], but had similar effects. Future work should examine whether other creative forms of non-realistic shadows [Jacquemin et al. 2011; Kasahara et al. 2019] would also match performance of realistic shadows in distance perception tasks.

Consistent with Gibson's ground theory, prior work has shown that AR objects are perceived as farther away when they are located off the ground, given no additional information for ground contact [Rosales et al. 2019]. However, when shadows are provided as a ground contact cue, we would expect this effect to be reduced. Thus, our current finding that there was no effect of height above the ground on perceived distance with both types of shadows provides further support that the stylized shadow serves as an effective cue for location, similar to the realistic shadow. We tested only one relatively small (20 cm) height above the ground. Future work could assess the utility of different types of shadows at different heights and distances to further generalize these effects.

This study used verbal reports as the response measure for distance. Distance underestimation is consistently seen in the real world when using verbal reports [Loomis and Philbeck 2008]. Because we did not run a matched real world condition, we cannot make strong claims about whether the current AR results show greater or comparable amounts of underestimation to the real world. However, it is notable that the magnitude of distance underestimation increased with increasing distance in our data (~20% at 3m, ~37% at 4.5m, and ~44% at 6m), a result that is not typically seen in the real world at this range of distances [Loomis and Philbeck 2008; Philbeck and Loomis 1997]. Studies that have directly compared AR and real world estimates show mixed results, which is likely due to multiple factors including the type of measure and the type of AR display. Action measures, such as blind walking [Loomis et al. 1992; Rieser et al. 1990] are typically accurate in the real world and some studies have shown comparable distance estimation with real and AR targets [Stefanucci et al. 2021]. There has been significant work in virtual reality studying distance estimation and response measures as well (see Creem-Regehr et al. [2023] for a recent review), but less work has been done in AR (however, see Jamiy et al. [2020] for a video see-through virtual reality comparison). It would be interesting to compare these AR devices in a study with action measures to see if it leads to an improved understanding of the differences between OST and VST devices.

6 CONCLUSIONS

This paper examined egocentric distance perception in two modern AR HMDs using verbal reports. We varied virtual objects contact with the ground and shadow type, varying the latter between a realistic and non-realistic shadow. Consistent with prior work, we found significant underestimation of distance in both displays, but less underestimation in the Microsoft Hololens 2 than the Varjo XR-3. We found little evidence for differences in the accuracy of distance judgments based on realistic or non-realistic shadow conditions, or on height above the ground, which was confirmed by a Bayesian analysis. Thus, non-realistic shadows may serve as good surface contact cues in some applications where realistic shadows could be problematic. Future work could investigate this further.

ACKNOWLEDGMENTS

The authors thank the reviewers for their helpful comments. This work was supported by National Science Foundation grants 1763254 and 1763966 and Office of Naval Research grants N00014-18-1-2964 and N0014-21-1-2583.

REFERENCES

Haley Adams, Jeanine Stefanucci, Sarah Creem-Regehr, and Bobby Bodenheimer. 2022. Depth perception in augmented reality: The effects of display, shadow, and position. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 792–801.

Haley Adams, Jeanine Stefanucci, Sarah Creem-Regehr, Grant Pointon, William Thompson, and Bobby Bodenheimer. 2021. Shedding light on cast shadows: An investigation of perceived ground contact in ar and vr. *IEEE Transactions on Visualization and Computer Graphics* 28, 12 (2021), 4624–4639.

Murray Aitkin. 1991. Posterior bayes factors. *Journal of the Royal Statistical Society: Series B (Methodological)* 53, 1 (1991), 111–128.

Jeffrey Andre and Sheena Rogers. 2006. Using verbal and blind-walking distance estimates to investigate the two visual systems hypothesis. *Perception & Psychophysics* 68, 3 (2006), 353–361.

D Bates, M Mächler, B Bolker, and S Walker. 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software,* 67, 1 (2015), 1–48.

O. J. Berger. 2006. Bayes Factors. In *Encyclopedia of Statistical Sciences*, S. Kotz, N. Balakrishnan, C. Read, B. Vidakovic, and N. L. Johnson (Eds.). Wiley, Hoboken, NJ, 378–386.

Lauren E. Buck, Mary K. Young, and Bobby Bodenheimer. 2018. A Comparison of Distance Estimation in HMD-Based Virtual Environments with Different HMD-Based Conditions. *ACM Trans. Appl. Percept.* 15, 3, Article 21 (July 2018), 15 pages. https://doi.org/10.1145/3196885

PC Bürkner. 2017. Bayesian Regression Models using Stan. *R package version* 1, 0 (2017).

Nadia Cattari, Fabrizio Cutolo, Renzo D'amato, Umberto Fontana, and Vincenzo Ferrari. 2019. Toed-in vs Parallel Displays in Video See-Through Head-Mounted Displays for Close-Up View. *IEEE Access* 7 (2019), 159698–159711. https://doi.org/10.1109/ACCESS.2019.2950877

Sarah H. Creem-Regehr, Jeanine K. Stefanucci, and Bobby Bodenheimer. 2023. Perceiving distance in virtual reality: theoretical insights from contemporary technologies. *Philosophical Transactions of the Royal Society B: Biological Sciences* 378, 1869 (2023), 20210456. https://doi.org/10.1098/rstb.2021.0456 arXiv:https://royalsocietypublishing.org/doi/pdf/10.1098/rstb.2021.0456

James E. Cutting and Peter M. Vishton. 1995. Perceiving Layout and Knowing Distance: The Integration, Relative Potency and Contextual Use of Different Information about Depth. In *Perception of Space and Motion,* William Epstein and Sheena Rogers (Eds.). Academic Press, New York, 69–117.

Catherine Diaz, Michael Walker, Danielle Albers Szafir, and Daniel Szafir. 2017. Designing for Depth Perceptions in Augmented Reality. In *2017 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE, 111–122.

Fatima El Jamiy, Ananth N Ramaseri Chandra, and Ronald Marsh. 2020. Distance accuracy of real environments in virtual reality head-mounted displays. In *2020 IEEE International Conference on Electro Information Technology (EIT)*. IEEE, 281–287.

F. El Jamiy and R. Marsh. 2019. Survey on depth perception in head mounted displays: distance estimation in virtual reality, augmented reality, and mixed reality. *IET Image Processing* 13, 5 (2019), 707–712.

Holly C Gagnon, Carlos Salas Rosales, Ryan Mileris, Jeanine K Stefanucci, Sarah H Creem-Regehr, and Robert E Bodenheimer. 2021a. Estimating distances in action space in augmented reality. *ACM Transactions on Applied Perception (TAP)* 18, 2 (2021), 1–16.

Holly C Gagnon, Yu Zhao, Matthew Richardson, Grant D Pointon, Jeanine Stefanucci, Sarah H Creem-Regehr, and Bobby Bodenheimer. 2021b. Gap affordance judgments in mixed reality: Testing the role of display weight and field of view. *Frontiers in Virtual Reality* 2 (2021), 22.

Yuan Gao, Etienne Peillard, Jean-Marie Normand, Guillaume Moreau, Yue Liu, and Yongtian Wang. 2019. 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).

Andrew Gelman, John B Carlin, Hal S Stern, David B Dunson, Aki Vehtari, and Donald B Rubin. 2013. *Bayesian data analysis.* CRC press.

James J Gibson. 1950. *The perception of the visual world.* Houghton Mifflin.

Peter Grünwald. 2000. Model selection based on minimum description length. *Journal of mathematical psychology* 44, 1 (2000), 133–152.

Robert T. Held and Martin S. Banks. 2008. Misperceptions in Stereoscopic Displays: A Vision Science Perspective. In *Proceedings of the 5th Symposium on Applied Perception in Graphics and Visualization* (Los Angeles, California) *(APGV '08)*. Association for Computing Machinery, New York, NY, USA, 23–32. https://doi.org/10.1145/1394281.1394285

Sei Ikeda, Yuto Kimura, Shinnosuke Manabe, Asako Kimura, and Fumihisa Shibata. 2020. Shadow induction on optical see-through head-mounted displays. *Computers & Graphics* 91 (2020), 141–152. https://doi.org/10.1016/j.cag.2020.07.003

Christian Jacquemin, Georges Gagneré, and Benoît Lahoz. 2011. Shedding Light on Shadow: Real-Time Interactive Artworks Based on Cast Shadows or Silhouettes. In *Proceedings of the 19th ACM International Conference on Multimedia* (Scottsdale, Arizona, USA) *(MM '11)*. Association for Computing Machinery, New York, NY, USA, 173–182. https://doi.org/10.1145/2072298.2072322

J. Adam Jones, J. Edward Swan, II, Gurjot Singh, Eric Kolstad, and Stephen R. Ellis. 2008. 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* (Los Angeles, California) *(APGV '08)*. ACM, New York, NY, USA, 9–14. https://doi.org/10.1145/1394281.1394283

J Adam Jones, J Edward Swan II, and Mark Bolas. 2013. Peripheral stimulation and its effect on perceived spatial scale in virtual environments. *IEEE transactions on visualization and computer graphics* 19, 4 (2013), 701–710.

Alexandra D Kaplan, Jessica Cruit, Mica Endsley, Suzanne M Beers, Ben D Sawyer, and Peter A Hancock. 2021. The effects of virtual reality, augmented reality, and mixed reality as training enhancement methods: A meta-analysis. *Human factors* 63, 4 (2021), 706–726.

Shunichi Kasahara, Satoru Higa, and Akihiro Komori. 2019. Fragment Shadow: Generating Fragmented Shadows with Multi-Projectors Geometry and Color Calibration. In *ACM SIGGRAPH 2019 Studio* (Los Angeles, California) *(SIGGRAPH '19)*. Association for Computing Machinery, New York, NY, USA, Article 6, 2 pages. https://doi.org/10.1145/3306306.3328003

Jonathan W. Kelly. 2022. Distance Perception in Virtual Reality: A Meta-Analysis of the Effect of Head-Mounted Display Characteristics. *IEEE Transactions on Visualization and Computer Graphics* (2022), 1–13. https://doi.org/10.1109/TVCG.2022.3196606

Kiyoshi Kiyokawa, Yoshinori Kurata, and Hiroyuki Ohno. 2001. An optical see-through display for mutual occlusion with a real-time stereovision system. *Computers & Graphics* 25, 5 (2001), 765–779. https://doi.org/10.1016/S0097-8493(01)00119-4 Mixed realities - beyond conventions.

Liisa Kuparinen, J II, Scott Rapson, and Christian Sandor. 2013. Depth perception in tablet-based augmented reality at medium-and far-field distances. In *SAP'13 Proceedings of the ACM Symposium on Applied Perception (pp. 121). New York: Association for Computing Machinery. ISBN 978-1-4503-2262-1.* Association for Computing Machinery.

Mikko Kytö, Aleksi Mäkinen, Timo Tossavainen, and Pirkko T. Oittinen. 2014. Stereoscopic depth perception in video see-through augmented reality within action space. *Journal of Electronic Imaging* 23, 1 (2014), 1 – 11. https://doi.org/10.1117/1.JEI.23.1.011006

Wenkai Li, A. Y. C. Nee, and S. K. Ong. 2017. A State-of-the-Art Review of Augmented Reality in Engineering Analysis and Simulation. *Multimodal Technologies and Interaction* 1, 3 (2017). https://doi.org/10.3390/mti1030017

Charles C Liu and Murray Aitkin. 2008. Bayes factors: Prior sensitivity and model generalizability. *Journal of Mathematical Psychology* 52, 6 (2008), 362–375.

Jack M. Loomis and John W. Philbeck. 2008. Measuring Spatial Perception with Spatial Updating and Action. In *Embodiment, Ego-Space, and Action*, R. L. Klatzky, B. MacWihnney, and M. Behrmann (Eds.). Taylor & Francis, New York.

Jack M. Loomis, José A. Da Silva, Naofumi Fujita, and Sergio S. Fukusima. 1992. Visual Space Perception and Visually Directed Action. *Journal of Experimental Psychology: Human Perception and Performance* 18, 4 (1992), 906–921.

Daniel Lüdecke. 2018. sjstats: Statistical functions for regression models. *R package version 0.14. 0* (2018).

Cindee Madison, William Thompson, Daniel Kersten, Peter Shirley, and Brian Smits. 2001. Use of interreflection and shadow for surface contact. *Perception & Psychophysics* 63, 2 (2001), 187–194.

Pascal Mamassian, David C Knill, and Daniel Kersten. 1998. The perception of cast shadows. *Trends in cognitive sciences* 2, 8 (1998), 288–295.

Sina Masnadi, Kevin Pfeil, Jose-Valentin T Sera-Josef, and Joseph LaViola. 2022. Effects of Field of View on Egocentric Distance Perception in Virtual Reality. In *CHI Conference on Human Factors in Computing Systems.* 1–10.

Sina Masnadi, Kevin P Pfeil, Jose-Valentin T Sera-Josef, and Joseph J LaViola. 2021. Field of view effect on distance perception in virtual reality. In *2021 IEEE Conference* *on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. IEEE, 542–543.

Rui Ni, Myron Braunstein, and George Andersen. 2005. Distance perception from motion parallax and ground contact. *Visual Cognition* 12, 6 (Aug. 2005), 1235–1254.

Kevin Pfeil, Sina Masnadi, Jacob Belga, Jose-Valentin T Sera-Josef, and Joseph LaViola. 2021. Distance perception with a video see-through head-mounted display. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–9.

John W Philbeck and Jack M Loomis. 1997. 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 (1997), 72.

Grant Pointon, Chelsey Thompson, Sarah Creem-Regehr, Jeanine Stefanucci, Miti Joshi, Richard Paris, and Bobby Bodenheimer. 2018a. Judging action capabilities in augmented reality. In *Proceedings of the 15th ACM Symposium on Applied Perception*. 1–8.

Grant Pointon, Chelsey Thompson, Sarah Creem-Regehr, Jeanine Stefanucci, Miti Joshi, Richard Paris, and Bobby Bodenheimer. 2018b. Judging Action Capabilities in Augmented Reality. In *Proceedings of the 15th ACM Symposium on Applied Perception* (Vancouver, British Columbia, Canada) *(SAP '18)*. ACM, Article 6, 8 pages. https://doi.org/10.1145/3225153.3225168

R Core Team. 2022. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Kristina M Rand, Margaret R Tarampi, Sarah H Creem-Regehr, and William B Thompson. 2011. The importance of a visual horizon for distance judgments under severely degraded vision. *Perception* 40, 2 (2011), 143–154.

Kristina M Rand, Margaret R Tarampi, Sarah H Creem-Regehr, and William B Thompson. 2012. The influence of ground contact and visible horizon on perception of distance and size under severely degraded vision. *Seeing and perceiving* 25, 5 (2012), 425–447.

John J. Rieser, Danial H. Ashmead, Charles R. Tayor, and Grant A. Youngquist. 1990. Visual perception and the guidance of locomotion without vision to previously seen targets. *Perception* 19 (1990), 675–689.

Carlos Salas Rosales, Grant Pointon, Haley Adams, Jeanine Stefanucci, Sarah Creem-Regehr, William B Thompson, and Bobby Bodenheimer. 2019. Distance Judgments to On-and Off-Ground Objects in Augmented Reality. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 237–243.

Paulo E. Santos, Roberto Casati, and Patrick Cavanagh. 2018. Perception, cognition and reasoning about shadows. *Spatial Cognition & Computation* 18, 2 (2018), 78–85. https://doi.org/10.1080/13875868.2017.1377204 arXiv:https://doi.org/10.1080/13875868.2017.1377204

development team STAN. 2018. RStan: The R interface to Stan. R package version 2.17. 3. *Online: http://mc-stan. org* (2018).

Jeanine K Stefanucci, Sarah Creem-Regehr, and Bobby Bodenheimer. 2021. Comparing Distance Judgments in Real and Augmented Reality. In *2021 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. IEEE, 82–86.

JE Swan, Mark A Livingston, Harvey S Smallman, Dennis Brown, Yohan Baillot, Joseph L Gabbard, and Deborah Hix. 2006. A perceptual matching technique for depth judgments in optical, see-through augmented reality. *Proceedings of the IEEE Virtual Reality Conference, 2006* (2006), 19–26.

J Edward Swan, Adam Jones, Eric Kolstad, Mark A Livingston, and Harvey S Smallman. 2007. Egocentric depth judgments in optical, see-through augmented reality. *IEEE transactions on visualization and computer graphics* 13, 3 (2007), 429–442.

A. Takagi, S. Yamazaki, Y. Saito, and N. Taniguchi. 2000. Development of a stereo video see-through HMD for AR systems. In *Proceedings IEEE and ACM International Symposium on Augmented Reality (ISAR 2000)*. 68–77. https://doi.org/10.1109/ISAR.2000.880925

Koorosh Vaziri, Maria Bondy, Amanda Bui, and Victoria Interrante. 2021. 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)*. IEEE, 1–9.

Koorosh Vaziri, Peng Liu, Sahar Aseeri, and Victoria Interrante. 2017. 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*. 1–8.

Andrew J. Woods, Tom Docherty, and Rolf Koch. 1993. Image distortions in stereoscopic video systems. In *Stereoscopic Displays and Applications IV*, John O. Merritt and Scott S. Fisher (Eds.), Vol. 1915. International Society for Optics and Photonics, SPIE, 36 – 48. https://doi.org/10.1117/12.157041