Measurement Protocols for Medium-Field Distance Perception in Large-Screen Immersive Displays*

Eric Klein[†]
NVIDIA
Corporation

J. Edward Swan II[‡]
Mississippi State
University

Gregory S. Schmidt[§]
SPADAC
Inc.

Mark A. Livingston[§]
Naval Research
Laboratory

Oliver G. Staadt[†]
University of
Rostock

ABSTRACT

How do users of virtual environments perceive virtual space? Many experiments have explored this question, but most of these have used head-mounted immersive displays. This paper reports an experiment that studied large-screen immersive displays at medium-field distances of 2 to 15 meters. The experiment measured egocentric depth judgments in a CAVE, a tiled display wall, and a real-world outdoor field as a control condition. We carefully modeled the outdoor field to make the three environments as similar as possible. Measuring egocentric depth judgments in large-screen immersive displays requires adapting new measurement protocols; the experiment used timed imagined walking, verbal estimation, and triangulated blind walking.

We found that depth judgments from timed imagined walking and verbal estimation were very similar in all three environments. However, triangulated blind walking was accurate only in the outdoor field; in the large-screen immersive displays it showed underestimation effects that were likely caused by insufficient physical space to perform the technique. These results suggest using timed imagined walking as a primary protocol for assessing depth perception in large-screen immersive displays. We also found that depth judgments in the CAVE were more accurate than in the tiled display wall, which suggests that the peripheral scenery offered by the CAVE is helpful when perceiving virtual space.

Keywords: Distance Perception, Egocentric Depth Perception, Virtual Environments, Large-Screen Immersive Displays

Index Terms: I.2.10 [Artifical Intelligence]: Vision and Scene Understanding—Perceptual Reasoning; H.5.2 [Information Interfaces and Presentation]: User Interfaces—Ergonomics

1 INTRODUCTION

How *egocentric depth perception* operates in virtual environments (VEs) at medium-field distances of about 2 to about 20 meters has been extensively studied for the past 10 to 15 years; both Loomis and Knapp [13] and Swan et al. [16] survey this literature. Egocentric depth perception is the perception of the distance from an observer to objects in the environment. This is an important area of study because (1) it is an interesting intellectual question in its own right, and (2) understanding how depth perception works is necessary to properly implement many VE applications. In particular, many of these studies have found that in VEs distances are compressed relative to the same distances in real-world settings; explaining this phenomena has motivated much work in the area.

1.1 Distance Perception Measurement Protocols

Because perception is an invisible cognitive state, depth perception studies must use some *measurement protocol* to obtain a *depth judgment* of a target object. A large number of different measurement protocols have been proposed and studied, each with advantages and disadvantages. This section surveys the most widely-used protocols; it is not a comprehensive listing. The surveyed protocols can be divided into the general categories of verbal estimation, visually guided actions, visually imagined actions, and perceptual matching.

In *verbal estimation* protocols (e.g., Loomis and Knapp [13]), the observer states the depth in terms of some familiar unit, such as feet, meters, etc., or as multiples of some extent that is visible in the scene. While verbal estimation reports are easy to collect, there is always the danger that observers are using cognitive knowledge that is unrelated to the perception of distance (Loomis and Knapp [13]). For example, if the target object is a chair, observers can use their knowledge of a chair's expected size to infer the distance, and experimenters can "fool" observers by using chairs that are larger or smaller than expected, but these cognitive issues can confound measuring a perception of distance.

A widely-used category of measurement protocols have been *visually guided actions* (e.g., Loomis and Knapp [13]), where participants view a target, and then without seeing the target undertake some bodily action, such as reaching, walking, or throwing, that indicates the distance to the target. A common visually guided action is *blind walking* (Figure 1), where the observer views the target object and then walks without vision to the remembered location of the object. With real world targets blind walking is very accurate out to at least 20 meters (Loomis and Knapp [13]). A related visually guided action is *triangulated blind walking* (Figure 2), where the observer views a target object, turns to face an oblique angle to the object, views the target again, and then covers their eyes and walks forward without vision. At some point the experimenter stops the observer, and still without vision the observer turns and faces the object. These actions describe one side and one angle of a triangle, where the side opposite the angle represents a depth judgment to the object. With real world targets triangulated blind walking is very accurate out to at least 15 meters (Loomis and Knapp [13]).

A distinct advantage of visually guided actions over verbal estimation is that the observer's perception of distance can be directly inferred from the action. A potential disadvantage is the danger that the action comes from the calibration of the human body to everyday perceptual motor activity, as opposed to a perception of distance. However, strong evidence against this calibration hypothesis comes from studies where the observer's response is indirectly

---

coupled to the target distance. For example, in triangulated blind walking observers are accurate even though the distance that they walk along the oblique angle is arbitrary and unpredictable, and it is unlikely that this action could be previously calibrated from normal perceptual motor activity (Loomis and Knapp [13]).

Closely related to visually guided actions are *visually imagined actions*, where the action is imagined instead of actually performed. In *timed imagined walking* (Figure 3), the observer views a target object, closes their eyes, and then imagines walking to the object. The time it takes them to imagine walking to the object is recorded, and then combined with their measured walking rate to yield a depth judgment. Since the observer is standing still, an advantage of this technique is that it doesn't require any space. A disadvantage is that the action is imagined instead of performed, which involves potentially confounding mental processes. However, Decety et al. [4] and Plumert et al. [14] compared blind walking and timed walking to real-world targets, and found excellent accuracy for both methods.

In *perceptual matching* protocols, the observer indicates the distance of the target object by manipulating or judging the distance to a matching object. The matching object can either be positioned to one side of the target object (e.g., Ellis and Menges [5]), or positioned at the same distance in a different direction than the target object (e.g., in Wu et al. [19] the observer positions the matching object in a direction offset 90° from the direction of the target object). In *perceptual bisection* (e.g., Lappin et al. [12]), the observer positions the matching object at the bisection (midpoint) of the distance to the target object. For all perceptual matching protocols, there are two different ways that observers can indicate the location of the matching object: (1) with the *method of adjustment*, observers physically adjust the position of the matching object, either through some mechanical linkage or by telling the experimenter where to place the matching object; (2) with the *method of constant stimuli*, observers view the matching object, and then judge whether it is closer or farther than either the target object itself (perceptual matching), or the midpoint to the target object (perceptual bisection).

One advantage of perceptual matching protocols is that they rely only on visual perception. Another advantage is that they seem related to many useful VE applications; one such application is modeling in augmented reality (Wither and Höllerer [18]). A particular disadvantage of perceptual bisection is that it does not give an absolute measurement of perceived distance.

1.2 Measurement Protocols in Virtual Environments

All of the measurement protocols described above have been used to measure depth judgments in virtual environments; again the listing in this paragraph is not comprehensive. By far the most commonly used protocol has been blind walking (e.g., [7, 8, 10, 11, 13, 15, 16, 17, 19]). However, blind walking requires a large amount of space: there must be a clear path to the target, and a substantial amount of clear space between the target and any solid object, such as a wall, with which the observer might collide if they overshoot the target. Triangulated blind walking has also been widely used in virtual environments (e.g., [10, 13, 15, 17]); Thompson et al. [17] cite the space requirements of blind walking as a motivation for using triangulated walking. Another shortcoming of blind walking, which relates to the experiment reported here, is that it cannot be used to indicate a depth judgment in a large-screen immersive display, because there is not enough room to blindly walk to a target that is located beyond the display's screen. Plumert et al. [14] and Ziemer et al. [20] used timed walking to measure distance judg-
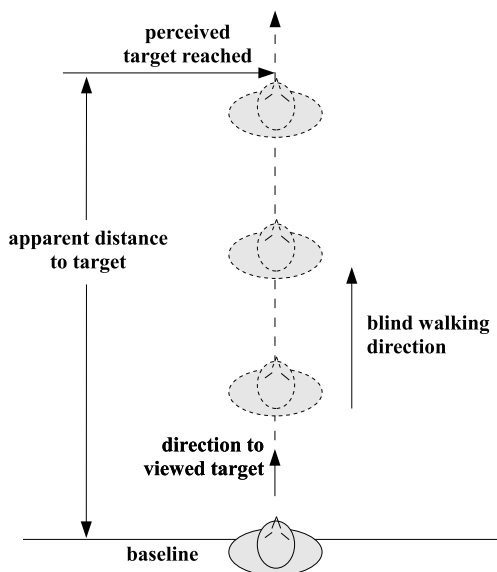

Figure 1: The *blind walking* depth judgment protocol. The observer views the target object, closes their eyes, and walks without vision to the remembered location of the target object.

ments in a CAVE. Perceptual matching has also been used in virtual environments (e.g., [5, 16, 18, 19]); Bodenheimer et al. [2] used perceptual bisection to measure distance judgments in virtual and real-world environments. Finally, verbal report has been used as well (e.g., [10, 11, 13, 16]).

In the current experiment, we wanted to compare several measurement protocols for obtaining depth judgments with large-screen immersive displays. In particular, we compared verbal estimation, timed imagined walking, and triangulated blind walking. All of these protocols can indicate larger depth judgments than the available physical space — verbal estimation and timed walking require no observer movement — and so all of them can potentially be used with large-screen immersive displays. Although timed walking has been previously studied with large-screen displays, we are only aware of the two published articles from the University of Iowa (Plumert et al. [14], Ziemer et al. [20]). Finally, to the best of our knowledge, triangulated walking has not yet been studied with large-screen displays. We decided to compare both protocols to verbal estimation because it represents a different measurement protocol category.

1.3 Distance Perception and VE Display Devices

The great majority of the previous virtual environment depth perception studies have used head-mounted, immersive display systems (e.g., [2, 7, 8, 10, 13, 15, 17, 18]). This leads to another motivation for the current experiment: many of these previous studies have carefully isolated experimental participants from obtaining spatial knowledge of the real-world location where the depth judgments are measured (e.g., Thompson et al. [17]). A typical protocol is for the experimenter to blindfold the participant, and then have the participant walk for approximately 10 minutes without vision under the experimenter's voice command. Next, the experimenter leads the participant into the room where the experiment will take place, and places the immersive head-mounted display on the participant's head while their eyes are closed and the room lights are off. Because of this protocol, the participant has no connection
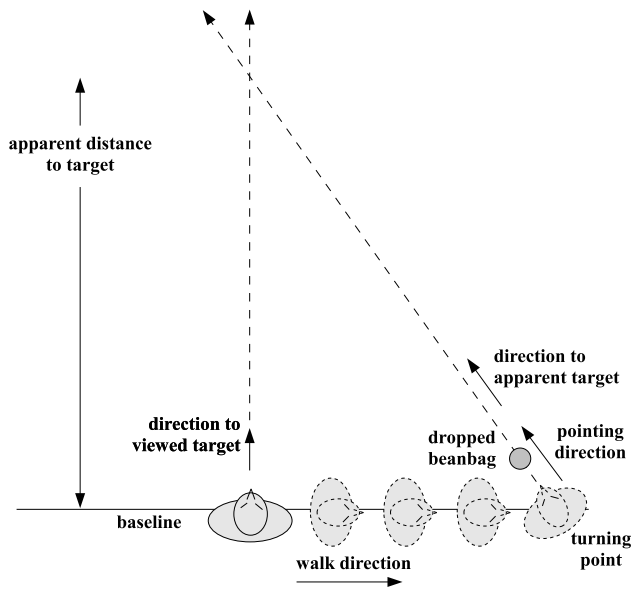
Figure 2: The *triangulated blind walking* depth judgment protocol. The observer views the target object, turns 90°, views the target again, closes their eyes, walks forward 2.5 meters, turns to face the object, points to the object with their outstretched hands, and then drops a beanbag held between their hands. These actions describe one side and one interior angle of a right triangle, where the opposite side represents a depth judgment of the target object.
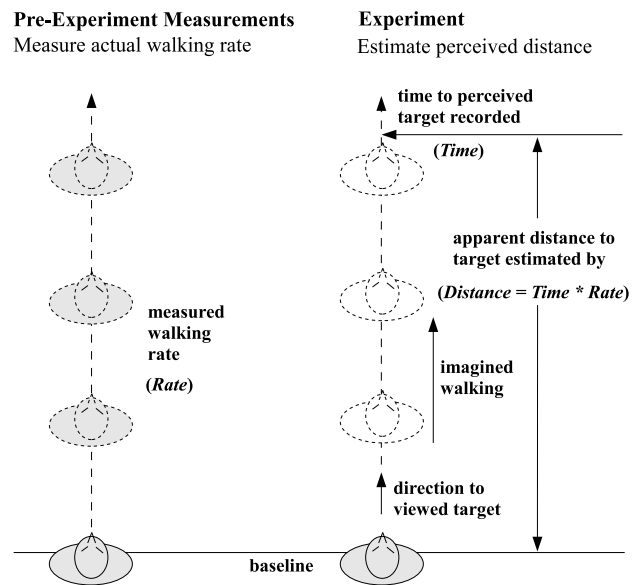

Figure 3: The *timed imagined walking* depth judgment protocol. Before collecting data, we measure each observer's walking speed. For each trial, the observer views the target object, closes their eyes, starts a timer, and imagines walking to the object. When they imagine reaching the object, the observer stops the timer and reports the elapsed time.

between the spatial layout of the virtual environment and the real-world location. It is believed that this isolation increases experimental validity; knowledge of the real-world location will make participants fearful that they may bump into a real-world obstacle or wall as they walk blindly. Interrante et al. [7] have found evidence that knowledge that an immersive environment is a faithful copy of a familiar real-world location can influence depth judgments.

However, there are virtual environment situations where observers are always going to be aware of both the VE and the real-world setting where the VE display device is located. These include both augmented reality (e.g., [5, 8, 16]) and large-screen immersive displays [14, 20]. It may be that depth perception operates differently when the virtual world exists as an extension of the real world, and this motivates experiments that examine these situations.

In the current experiment, we wanted to obtain depth judgments in large-screen immersive displays. There were two such displays on the University of California, Davis campus: a 4-wall CAVE and a tiled wall (Figure 4). The primary difference between the displays is the peripheral vision provided by the wrap-around walls and floor of the CAVE relative to the wall; other technical differences between the displays are described in Section 3.1 below. We also included a real-world open field on the campus as a control condition (Figure 5). We carefully modeled the outdoor field in the virtual environments to make the three environments as similar as possible. To the best of our knowledge, this is the first time VE depth perception has been studied with a wall display, and it is the first time two different large-screen immersive displays have been compared with each other in the same experimental context.

2 PREVIOUS WORK

To the best of our knowledge, the only previous distance perception studies with a large-screen immersive display have been con-

ducted at the University of Iowa. Plumert et al. [14] conducted the studies in a 3-walled CAVE, with a front wall and two side walls; CAVE graphics were presented non-stereoscopically, but were viewed binocularly. The CAVE was one environment; the other was a grassy field in front of a large building, which was modeled and displayed in the CAVE. Participants utilized a timed imagined walking procedure, but unlike Decety et al. [4] participants kept their eyes open as they imagined walking to the target. Experiment I found an effect of presentation order: participants who experienced the real environment first underestimated distance less than participants who experienced the virtual environment first; otherwise performance was very similar in the two environments. Experiment II found an effect of age: 10-year-olds demonstrated more underestimation in the virtual than in the real environment, but 12-year-olds and adults performed similarly. Experiment II also studied environment presentation order and whether participants were sighted or blind while performing timed walking, but did not find effects of either variable. Experiment III again studied sighted versus blind timed walking, and compared it to standard blind walking, in the outdoor environment only; this experiment found close agreement between timed walking and blind walking, and no effect of sighted versus blind timed walking. Ziemer et al. [20] report two follow-on studies using the same setup. Experiment IV again examined order effects, and replicated the presentation order effect of Experiment I. Experiment V found that the benefit of experiencing the real world before the virtual world was robust even when the particular outdoor location changed between the real and the virtual world. Overall, this series of experiments demonstrates (1) a close agreement between depth perception in the virtual and real worlds, (2) that it is not necessary for participants to be blind when they use the timed walking protocol, and (3) that real versus virtual world presentation order can make a difference in depth judgments.

Figure 4: A participant observing the field model on the Wall display. Also visible is the target object (a brightly-wrapped present).

3 METHOD

As stated above, our goal was to study depth judgments in three different environments (real world, CAVE, wall), using three different measurement protocols (timed imagined walking, verbal estimation, triangulated blind walking). We studied medium-field distances from 2 to 15 meters. Table 1 describes the basic experimental design.

3.1 Experimental Setup

We tested participants in three environments: a *Real-world* outdoor environment consisting of an open, grassy field (Figure 5); a four-wall stereoscopic *CAVE*; and a large, stereoscopic tiled *Wall* (Figure 4). We selected a flat, open field to maximize open space while minimizing visual interruptions such as shadows or patchy grass. We created a virtual environment model which mimicked this outdoor environment as closely as possible. This included using a photographic panorama background for distant objects, and a realistic Wang-tiled grass plane.

The Wall environment (Figure 4) consisted of a large continuous screen (5.49 meters across by 2.74 meters high) composed of six tiles (three across and two down), each 1024×768 in resolution. Each tile was powered by two projectors which were shuttered at 120 Hz to provide a stereo image when the participant was equipped with a pair of shutter glasses. An inertial/ultrasonic tracking system (with the tracker attached to the shutter glasses) was used to provide head tracking so that the user could make small head and body movements and see the results rendered realistically. Participants made observations from a position 1.22 meters back from the center of the screen.

The CAVE environment consisted of three walls and a floor. The CAVE measured 3.05 meters deep by 3.05 meters across by 2.44 meters tall. Each surface was 1400×1050 in resolution and images were displayed at 120 Hz, alternating between left and right images. As with the Wall environment, head tracking was provided by an inertial/ultrasound system, and the participant wore shutter glasses to perceive the stereo image. The participant made observations along the centerline of the CAVE from 0.91 meters outside of the CAVE entrance (to allow for triangulated blind walking as discussed below).

Figure 5: The field where the real-world outdoor testing occurred.

3.2 Participants

We recruited 23 participants from the university community of Davis, California; 14 were male and 9 were female. We screened the participants for 20/20 vision, either natural or corrected, out of both eyes. All participants volunteered, and were not compensated for their participation. As described in Section 4 below, we only retained the data from 20 participants for analysis.

3.3 Experimental Task

Before collecting data, we timed each participant as they walked 16.15 meters (53 feet); we repeated this 3 times and computed their average walking rate. For all trials, we first asked participants to close their eyes. While their eyes were closed, during real-world trials we walked to the farthest marker position (15 meters) and back, placing the test object in the required position along the way; this took 20–30 seconds. During virtual-world trials we walked to the computer terminal and pressed a key to make the test object appear; this took 5–10 seconds. We next asked the subject to open their eyes and judge the distance to the test object. We allowed participants to observe the object for as long as desired before making their depth judgment.

For a *timed imagined walking* judgment (Figure 3), we asked the participant to close their eyes and visualize walking to the object that they saw. We instructed them to start a stopwatch when they started walking, and stop it when they arrived at the object. We recorded this time, and used the participant's measured walking rate to compute a depth judgment. For a *verbal estimation* judgment, we asked the participant to state the distance to the object in whatever units they were most comfortable using (18 participants used feet, 2 used meters, and 3 used yards). For a *triangulated blind walking* judgment (Figure 2), the participant held a spherical beanbag as they observed the object. When the participant was ready to make a judgment, we asked them to turn 90° to the right, observe the object one more time, look forward, and close their eyes. With eyes closed, the participant walked until we instructed them to stop. Due to space constraints in the virtual displays, this stopping point was approximately 2.5 meters from the origin; we used the same distance in the real world setting. After stopping, while keeping their eyes closed, we asked the participant to turn and face the object, stand straight, and hold both hands flat together out in front of them, with the beanbag held between their hands, pointing at the distant object. We then asked them to drop the beanbag, and we made two marks for later recording, one for a position between their heels, and another for the position of the dropped beanbag. We placed a

Table 1: Independent and Dependent Variables

INDEPENDENT VARIABLES		
participant	20	(random variable)
environment	3	Real world
		CAVE
		Wall
protocol	3	Timed imagined walking
		Verbal estimation
		Triangulated blind walking
distance	6	2, 3, 6, 10, 15 meters
repetition	2	1, 2
DEPENDENT VARIABLES		
judged distance, in meters		
$normalized\ distance = \frac{judged\ distance}{actual\ distance} \times 100\%$		

numbered marker next to each mark (indoor markers were circular sticky dots, outdoor markers were golf tees). After the participant completed the experiment, we carefully measured the position of each numbered marker; during preliminary testing we found it took much too long to make these measurements in between participant trials.

We based this triangulated walking protocol on the one previously used by Loomis and Knapp [10, 11, 13] in both real-world and virtual environments (viewed with a head-mounted display). Ideally, we would have had participants turn an acute angle less than $90°$, walk more than 2.5 meters from the origin, and then after stopping and facing the object, walk a few paces in the direction of the object. However, we had to adopt the protocol to work in the rooms where the CAVE and Wall are located. The 2.5 meter walking distance ensured that participants never approached a physical object, such as a wall or a desk, closer than 2 meters.

For the target object we used a 39 cm $\times$ 26 cm $\times$ 30 cm box (Figure 4), which we wrapped in brightly-colored reddish-purple wrapping paper that provided good contrast with the saturated green grass. In the virtual environment we modeled the box with photographic textures and shadows to mimic the real box as closely as possible. We placed the box 2, 3, 6, 10, 12, and 15 meters from the participants, replicating the distances tested by Knapp and Loomis [11]. We tested two repetitions of every combination of the independent variables.

3.4 Dependent Variables

Our primary dependent variable was *judged distance* (Table 1). In addition, we calculated *normalized distance*; a normalized distance near 100% is veridical, while a normalized distance $> 100\%$ indicates overestimation, and a normalized distance $< 100\%$ indicates underestimation.

3.5 Experimental Design

We used a factorial nesting of independent variables in this within-subjects design, which varied in the order that they are listed in Table 1. *Environment* varied the slowest; within each environment participants made depth judgments with each *protocol*. The presentation order of environments and protocols was counterbalanced with nested between-subjects 3×3 Latin Squares. Within each *environment $\otimes$ protocol* block, we generated a list of 6 (*distance*) $\times$ 2 (*repetition*) = 12 distances, and then randomly permuted the pre-

sentation order, with the restriction that the same distance could not be presented twice in a row.

This design has the properties that environment presentation order is counterbalanced modulo 3 participants, and protocol presentation order is counterbalanced modulo 9 participants. In addition, environment succeeding and preceding order[1] is counterbalanced modulo 18 participants, and protocol succeeding and preceding order is counterbalanced modulo 36 participants. These properties counterbalance presentation order effects (to some degree of power), such as the real versus virtual world effects found by Plumert et al. [14] and Ziemer et al. [20].

4 DATA COLLECTION AND PROCESSING

Our goal was to collect data from a perfectly-counterbalanced set of 18 participants. However, the logistics proved extremely challenging. We could only collect real-world data during nice weather, but we collected data during the rainy season in Northern California. Often, when scheduled participants arrived, we could not collect data in the desired environmental order, and some participants had to return during subsequent days to complete data collection. It took collecting data from a total of 23 participants to obtain a perfectly-counterbalanced subset of 18 participants. Our design allowed us to examine this 18-participant subset for presentation order effects (again, to a certain degree of power). However, we could not find any systematic order effects in this 18-participant subset.

Because the data lacked order effects, and because the data from all 23 participants contained more power, we analyzed the full 23-participant dataset. This full dataset consisted of 2484 data points. It contained both missing data points and outliers, which we processed using techniques described by Barnett & Lewis [1] and Cohen et al. [3]. 15 data points were missing, and represented data entry errors (primarily from the triangulated walking trials, when a data marker could not be found). We judged 12 data points to be outliers; these included negative distances from triangulated walking trials with indicated angles $> 90°$. We replaced these 27 data points using either the remaining value in the experimental cell, or (if both values were missing) by linearly interpolating from neighboring cells.

We next examined the data for each participant. We did not find significant participant differences for the environment condition, nor did we find significant participant differences for the triangulated walking and timed walking protocols. However, for the verbal estimation protocol we found that three participants greatly overestimated the depth. For these overestimating participants normalized verbal distance = 203.3%, while for the remaining 20 participants normalized verbal distance = 75.0%, a difference of $d = 43.6$ standard errors. Furthermore, if we divide the $N = 828$ normalized verbal distances into an overestimating group (the three participants) and a non-overestimating group (the remaining 20 participants), a regression on these groups accounts for $r^2 = 23.3\%$ of the observed variance, and a discriminate analysis places 92.9% of the distances into the correct group. Including these three participants' data in the analysis significantly increases the verbal distance results. For these reasons, we eliminated these participants from further analysis. Therefore, Table 1 lists 20 experimental participants, and the next section presents the results from these 20 participants.

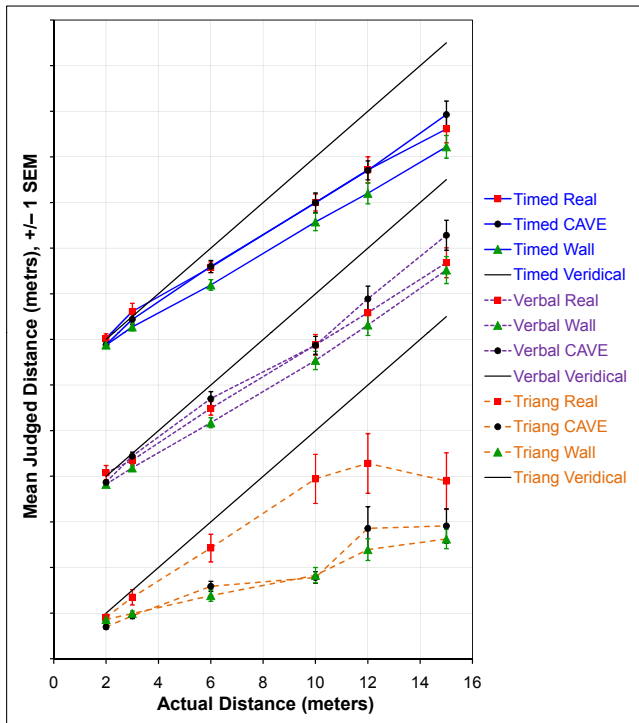[1]*Succeeding and preceding order* is described in Jones et al. [8].

Figure 6: Judged distance by actual distance ($N = 2160$). The diagonal lines are veridical; the results are offset by protocol for clarity.

5 RESULTS

Figure 6 shows the main results as a scatterplot between the mean judged distances and the actual distances. For clarity, the results are offset by protocol; the diagonal lines represent veridical performance. Figure 7 shows the same means as a scatterplot that is not offset by protocol; here the means are fitted with regression lines. The slopes of these lines give another estimate of the overall normalized distance for each protocol. Figure 8 shows normalized distances over environment and protocol; these are averaged over distance and repetition. A 3×3 repeated-measures ANOVA (using the Huynh-Feldt correction for non-sphericity (Howell [6])) on normalized distance gives a main effect of environment ($F(1.8, 33.6) = 11.68, p < .000, \varepsilon = .883$), a main effect of protocol ($F(2.0, 38.0) = 11.30, p < .000, \varepsilon = 1.000$), and an environment $\times$ protocol interaction ($F(2.2, 41.8) = 2.99, p = .056, \varepsilon = .550$). A post-hoc homogenous subset test (Howell [6]) on the 9 means in Figure 8 yields the homogenous groups indicated by the letters.

The most notable finding is that verbal estimation, timed walking, and real-world triangulated walking (through 12 meters) gave very similar results. This can be seen visually in Figure 6, in the remarkable overlap between the three regression lines in Figure 7, and in the homogenous subsets $A, B,$ and C in Figure 8. Although real-world triangulated walking showed a non-linear underestimation at 15 meters, when this point is excluded regression lines fit the remaining data very well (r^2 values of 98.5%, 98.5%, and 99.2%). The slopes of these lines (70.6%, 74.1%, 69.4%) indicate a general trend of underestimation for distances greater than 3 meters; the average value from Figure 8 is 81.8%. This degree of distance underestimation is comparable to what has been found for visually immersive VE displays (e.g., Table 2 in Thompson et al. [17] lists values from 44%–85%), although others have found normalized
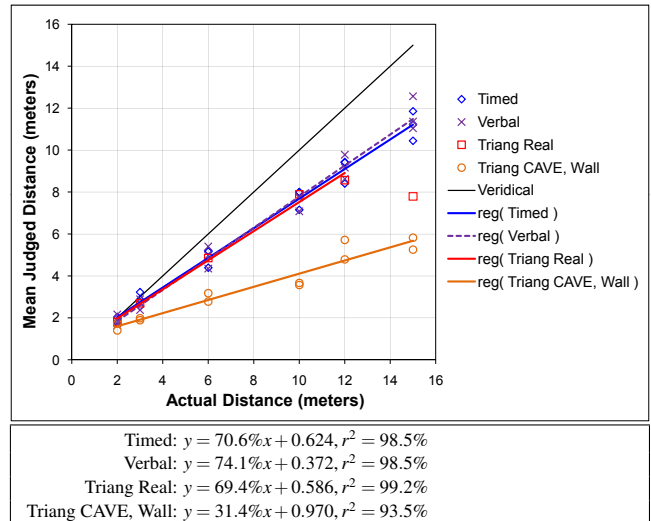

Timed: $y = 70.6\%x + 0.624, r^2 = 98.5\%$
Verbal: $y = 74.1\%x + 0.372, r^2 = 98.5\%$
Triang Real: $y = 69.4\%x + 0.586, r^2 = 99.2\%$
Triang CAVE, Wall: $y = 31.4\%x + 0.970, r^2 = 93.5\%$

Figure 7: Regression analysis of the means shown in Figure 6 ($N = 54$). The Triangulated Real regression line does not include the last point at 15 meters. Note that the slopes of the regression lines are not geometric slopes, but are percentages of the $45°$ veridical line.
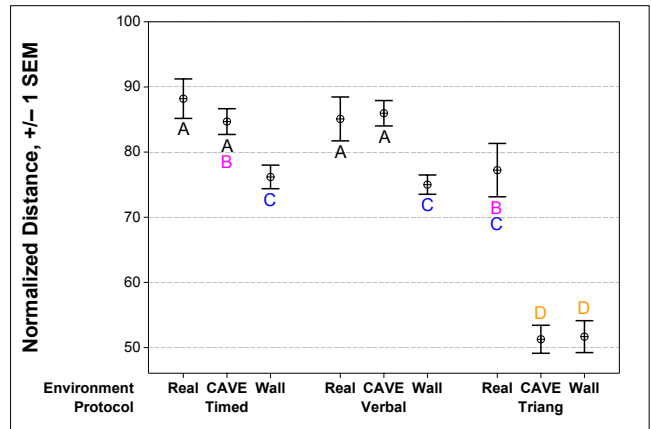

Figure 8: Normalized distance ($N = 2160$). Means with the same letter are not significantly different at $p \leq 0.05$ (Ryan REGWQ post-hoc homogenous subset test (Howell [6])).

distances from triangulated walking of 100% in the real world (e.g., Knapp [10]) and 91% in immersive VE displays (e.g., Richardson and Waller [15]).

Relative to these findings, triangulated walking underestimated all distances for the Wall and CAVE environments; the slope of the regression line in Figure 7 is 31.4%, while the average from Figure 8 is 51.5%. Figure 8 further suggests that this underestimation causes the majority of the ANOVA main effects and interaction; here the mean normalized distance for the subsets $\{A, B, C\}$ is 81.8% and for subset D is 51.5%, a difference of $d = 26.2$ standard errors.

The rest of the ANOVA effects come from a tendency to underestimate distances in the Wall relative to the other environments, especially for the timed walking and verbal estimation protocols (Figures 6 and 8). In particular, in Figure 8 note the relationship between subsets A and C for these two protocols: the mean normalized distance for subset A is 86.0% and for subset C is 76.2%, a difference of $d = 6.9$ standard errors.

6 CONCLUSIONS

We found a strong agreement between timed walking, verbal estimation, and triangulated walking in the real world, and a strong agreement between timed walking and verbal estimation in the CAVE and Wall displays. Similarly, Plumert et al. [14] and Ziemer et al. [20] found a strong agreement between real and virtual world performance with timed walking. Furthermore, our CAVE setup differed substantially from theirs: our 4-wall CAVE had a floor and presented the environment stereoscopically, while their 3-wall CAVE lacked a floor and presented the environment non-stereoscopically. These differences mean that timed walking has performed well on a variety of large-screen immersive displays, as well as outdoors. Taken together, the evidence supports using timed walking as a distance perception measurement protocol in large-screen immersive display systems, and perhaps more generally as well.

In contrast, triangulated walking did not work well in the CAVE or Wall. However, it did work well in the real-world environment (for distances up to 12 meters), which argues that we correctly implemented the basic triangulated walking technique: the 90° turn, the 2.5-meter baseline walk, the turn to face the target, the hand pointing, and then the beanbag drop worked well outdoors. This suggests that the problem indoors was participants' proximity to walls and obstacles in the room, and suggests that 2 meters of clearance is not enough to prevent the participants' knowledge of the room geometry from interfering with the triangulated walking task.

While we did not find the same presentation order effects as Plumert et al. [14] and Ziemer et al. [20]'s Experiments I and IV, their experiments were directly structured to study order effects, and so had more power to detect them. Similarly, Plumert et al.'s Experiment II did not replicate the order effects even though presentation order was included as an independent variable.

We also found evidence that distance perception is more accurate in a CAVE than a tiled Wall. This suggests that the peripheral scenery available in the CAVE may be helpful when perceiving virtual environment scale at medium-field distances; Plumert et al. [14] make a similar argument when discussing their results.

Finally, our results add to the great diversity of depth perception results that have been reported, both within real and virtual environments. As this paper has demonstrated, depth judgments have been collected with (1) a large number of different measurement protocols, in (2) a variety of outdoor and indoor settings, which have been viewed (3) both in the real world and in a variety of different VE display devices. Consider that recently Lappin et al. [12] found reliable, reproducible real-world depth judgment differences just by altering the environmental setting between an open field, a large room, and a hallway. Given these results, it seems clear that depth perception is influenced by many subtle aspects of the setting itself and how the setting is displayed to the observer. This calls for additional studies that carefully compare the many available parameters against each other.

REFERENCES

[1] V. Barnett and T. Lewis. *Outliers in Statistical Data*. Wiley Series in Probability & Statistics. John Wiley & Sons, 3rd edition, 1994.

[2] B. Bodenheimer, J. Meng, J. Wu, G. Narasimham, B. Rump, T. P. McNamara, T. H. Carr, and J. J. Rieser. Distance estimation in virtual and real environments using bisection. In *Proceedings of Symposium on Applied Perception in Graphics and Visualization (APGV 2007)*, pages 35–40, Tübingen, Germany, July 2007.

[3] J. Cohen, P. Cohen, S. G. West, and L. S. Aiken. *Applied Multiple Regression / Correlation Analysis for the Behavioral Sciences*. LEA, Inc., 2nd edition, 2002.

[4] J. Decety, M. Jeannerod, and C. Prablanc. The timing of mentally represented actions. *Behavioural Brain Research*, 34:35–42, 1989.

[5] S. R. Ellis and B. M. Menges. Localization of virtual objects in the near visual field. *Human Factors*, 40(3):415–431, 1998.

[6] D. C. Howell. *Statistical Methods for Psychology*. Duxbury, Pacific Grove, CA, 5th edition, 2002.

[7] V. Interrante, B. Ries, J. Lindquist, M. Keading, and L. Anderson. Elucidating factors that can facilitate veridical spatial perception in immersive virtual environments. *Presence: Teleoperators and Virtual Environments*, 17(2):176–198, 2008.

[8] J. A. Jones, J. E. Swan II, 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 Symposium on Applied Perception in Graphics and Visualization (APGV 2008)*, pages 9–14, Los Angeles, CA, USA, Aug. 2008.

[9] E. Klein, O. Staadt, J. E. Swan II, G. S. Schmidt, and M. A. Livingston. Egocentric medium-field distance perception in projection environments. In *Proceedings of Symposium on Applied Perception in Graphics and Visualization (APGV 2006)*, page 147, Boston, MA, USA, July 2006.

[10] J. M. Knapp. *The Visual Perception of Egocentric Distance in Virtual Environments*. PhD thesis, University of California, Santa Barbara, 1999.

[11] J. M. Knapp and J. M. Loomis. Limited field of view of head-mounted displays is not the cause of distance underestimation in virtual environments. *Presence: Teleoperators and Virtual Environments*, 13(5):572–577, 2004.

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

[13] J. M. Loomis and J. M. Knapp. Visual perception of egocentric distance in real and virtual environments. In L. J. Hettinger and J. W. Haas, editors, *Virtual and Adaptive Environments: Applications, Implications, and Human Performance Issues*, pages 21–46. Lawrence Erlbaum Associates, Mahwah, NJ, USA, 2003.

[14] J. M. Plumert, J. K. Kearney, J. F. Cremer, and K. Recker. Distance perception in real and virtual environments. *ACM Transactions on Applied Perception*, 2(3):216–233, 2005.

[15] A. R. Richardson and D. Waller. Interaction with an immersive virtual environment corrects users' distance estimates. *Human Factors*, 49(3):507–517, 2007.

[16] J. E. Swan II, 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 (TVCG)*, 13(3):429–442, 2007.

[17] W. B. Thompson, P. Willemsen, A. A. Gooch, S. H. Creem-Regehr, J. M. Loomis, and A. C. Beall. Does the quality of the computer graphics matter when judging distances in visually immersive environments? *Presence: Teleoperators and Virtual Environments*, 13(5):560–571, 2004.

[18] J. Wither and T. Höllerer. Pictorial depth cues for outdoor augmented reality. In *Proceedings of Ninth IEEE International Symposium on Wearable Computers (ISWC 2005)*, pages 92–99, Osaka, Japan, Oct. 2005.

[19] B. Wu, T. L. Ooi, and Z. J. He. Perceiving distance accurately by a directional process of integrating ground information. *Nature*, 428:73–77, 2004.

[20] C. Ziemer, J. M. Plumert, J. F. Cremer, and J. K. Kearney. Making distance judgments in real and virtual environments: Does order make a difference? In *Proceedings of Symposium on Applied Perception in Graphics and Visualization (APGV 2006)*, page 153, Boston, MA, USA, July 2006.