Evaluating Perceived Distance Measures in Room-Scale Spaces using Consumer-Grade Head Mounted Displays

Alex Peer* Kevin Ponto†

Living Environments Lab, Wisconsin Institute for Discovery
University of Wisconsin-Madison, USA

Abstract

Distance misperception (sometimes, distance compression) in immersive virtual environments is an active area of study, and the recent availability of consumer-grade display and tracking technologies raises new questions. This work explores the plausibility of measuring misperceptions within the small tracking volumes of consumer-grade technology, whether measures practical within this space are directly comparable, and if contemporary displays induce distance misperceptions.

Index Terms: I.3.7 [Computing Methodologies]: Graphics Utilities—Virtual reality

1 Introduction

Recent advancements in consumer-grade virtual reality devices are poised to share the potential of virtual reality (VR) with the world at large – but also its limitations. Discrepancies between the intended and perceived VR experience may affect task performance, quality of experience, and acceptance of VR. It is important that we study any potential discrepancies, and engineer corrections when possible; it is now becoming important that these corrective techniques be accessible to end-users of consumer-grade hardware. A well established discrepancy is that of distances being misperceived in immersive virtual environments. This effect is called distance compression, as the misperceptions commonly manifest as an underestimation. Renner et al. contributed a survey of the research on distance compression, finding that distances were estimated at 74% of their intended size in VR environments [17].

In this work, we evaluate distance compression measures that can be performed within the relatively small (4m x 4m) space tracked by a consumer-grade VR device. If established measures and consumer hardware are compatible, consumers should be able to measure, and possibly correct for, the degree of distance compression they are experiencing.

There is also some question as to whether improvements made to contemporary VR display systems have eliminated distance compression: some recent studies suggest so [3, 10, 11, 25], though earlier papers suggest the most obvious improvements, such as larger field of view, should not have so strong an effect [5,23]. To this end, we test two different contemporary head-mounted displays (HMDs) to see if either greatly reduces or eliminates distance compression.

Only a few studies have directly compared different perceived distance measures using the same visual stimulus and presentation context, generally establishing performance relative to blind walking [5, 13, 15, 19], which is impractical in small spaces. Only rarely have multiple measures been compared within a single work [8]. This makes it difficult to make comparisons across papers, which use varying measures, visual stimuli, and otherwise differing perceptual contexts. We use a fully-within participant design, observing each participant's performance with all measurement techniques, under all viewing conditions, using the same virtual and real environments. This provides a unified perceptual context for comparisons between measures and devices, as well as greater statistical power.

*e-mail: alex.peer@wisc.edu
†e-mail: kbponto@wisc.edu

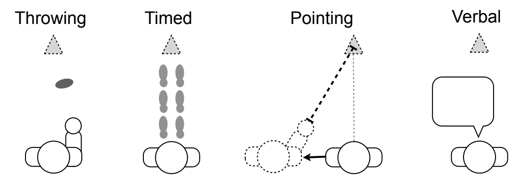

Figure 1: The four measures of distance perception used in this experiment: blind throwing, time-imagined walking, blind triangulated pointing, and verbal report.

2 Experiment

Virtual scenes were displayed using two consumer-grade head-mounted displays (HMDs): an Oculus Rift CV1, and an HTC Vive. According to manufacturer specifications, these HMDs have similar display characteristics: OLED screens with 2160x1200 resolutions and 90Hz refresh rates, and a 100 degree field of view. They vary slightly in weight, with the Vive weighing slightly more (555g) than the Rift (470g). During the experiment, both HMDs were adjusted for an interpupillary distance of 63.5mm. Rendering for both HMDs was provided through the Unity game engine, using OpenVR. Positional tracking was provided by each headset's respective tracking solution; the Vive provided tracking for two hand-held controllers, used by participants and experimenters to register participants' distance judgements.

The virtual environment (VE) displayed in the HMDs was made to be a rough, non-photorealistic match of the dimensions, color, and visual texture of the real environment (RE). We choose not to use a photorealistic match to avoid the previously studied effects on perceived distance of presentation order [26] and transitional environments [20]. 17 Participants were recruited from a local university campus, 11 men and 6 women ranging in age from 20 to 64 ($M = 27.6, SD = 10.1$).

2.1 Perceived Distance Measures

Blind Walking, in which a participant wearing an HMD walks to the perceived position of a previously viewed target, is one of the best established methods of measuring perceived distance [17], but requires significant space. In this experiment, we explore four other established measures designed for small spaces, which are depicted in Figure 1 and further described in this section.

Blind Throwing, as implemented here, follows the methods proposed by [14, 19]. After viewing the target, participants closed their eyes and threw a beanbag at the target. Participants were instructed to aim the center of the beanbag at the center of the target, and that the beanbag's initial point of impact would be recorded. Perceived distances were recorded by placing a tracked controller above the beanbag's point of impact and pushing a button to record its position, projected onto the floor plane, and on the axis running between the participant and target. Pilot trials suggested that participants would be hesitant to throw beanbags blind without practice, so they were allowed training throws at up to three target distances displayed in the real environment; training was terminated when participants felt comfortable with the task. Training distances were at intervals not seen in evaluated trials.

Timed Imagined Walking [5, 15, 26] asks participants to judge the amount of time they imagine it would take to walk to the target, using their separately measured average walking speed to convert this into a distance. In our experiment's implementation, the participant was first taken to a nearby hallway and asked to walk between two lines of tape placed 8 meters apart while the experimenter timed them with a stopwatch. This was repeated twice, and their average time to walk 8 meters was used to calculate their average walking speed. Participants then returned to the main experimental space and resumed trials. When performing the measurement task, participants held a tracked controller, viewed the target, then closed their eyes and imagined walking to the target's location. Participants pressed a button when they began their imagined walk, and again when they imagined they had reached the target's location; the first press began a timer, and the second stopped the timer and recorded the result. Perceived distance to the target was calculated by multiplying the participant's average walking speed, in meters per second, and their imagined walk duration, in seconds.

Blind Triangulated Pointing is a perceived distance measure used by [2], and is a space-bound adaption of the triangulated blind walking technique used by many [8, 18, 21, 22], and the several techniques presented in [4]. When performing this measurement task, participants hold a controller whose position and orientation are tracked. After viewing the target, participants close their eyes, take two steps to their left, then point a tracked controller toward the target and push a button to record the measurement. Before beginning the pointing phase of the experiment, participants viewed a demonstration by the experimenter and were asked to demonstrate themselves. During this demonstration, participants viewed a real-world target at a single distance interval different from those seen in the evaluated portion of the experiment. After recording their measurement, participants were prompted to return to their original position. Perceived distance was calculated by casting a ray from the position of the held controller in the direction the controller was pointed; the intersection between this ray, projected onto the ground plane, and the axis running between the participant and target was taken to be the perceived position of the target. This technique ignores any error in the vertical angle of the participant's pointing.

Verbal Report is a method of perceived distance measurement that has seen several variations in the literature [8, 9, 12, 16]. Our implementation is most similar to [8], as we simply asked participants to close their eyes and tell us how far away the target seemed, using whatever unit of measure they are most comfortable with (12 used feet, 5 used meters, and 1 used centimeters).

2.2 Procedure

Virtual environment trials progressed as follows: Participants begin with their eyes closed. By pushing a button they trigger a chime prompting them to open their eyes, as well as show the scene containing the target on the display device's screen; after three seconds the scene disappears, an audio cue plays to prompt participants to close their eyes, and they then perform the measurement task. Physical environment trials progressed similarly, but, as the experimenter manually placed the target between trials, participants waited to view the scene until prompted. The experimenter registered the position of the target by positioning the controller above it and pressing a button, then the moved away from the target and prompted the participant to pull the trigger and open their eyes. Three seconds later an audio cue prompted the participant to close their eyes, and they performed the measurement task. Participants triggered trial progression themselves in all conditions but the blind throwing measure, where they held beanbags rather than a controller; in this case, the experimenter triggered the next trial, after verbally warning the participant.

Before the experiment began, participants read and signed a consent form. They then were introduced to the audio cues meant as prompts for opening and closing their eyes. Next, they wore and adjusted both headsets for clear viewing, and for comfortable wear and removal. They were then introduced to the controller, and shown how to operate the trigger. The general form of the experiment was then described – pull trigger, view a scene, perform an action. The first measurement task was then introduced. When switching to a new measurement task, an explanation of the task was given (see: §2.1), and it was confirmed that participants understood the task.

2.3 Methods

The experiment follows a within subjects, repeated-measures design, with all participants experiencing all conditions.

The independent variables are the perceived distance measurement task (*measure*), display device (*display*), and target distance (*distance*). *Measure* is a factor of four levels, as described in §2.1: (*throwing*, *pointing*, *walking*, *verbal*). *Display* is a factor of three levels: the RE (*real*), and the VE in both HMDs (*rift*, *vive*). Distance is a factor of three levels, corresponding to the three base distances used in our experiment: (*2m*, *3m*, *4m*).

The dependent variable is *percent error* (PE) of measured perceived distance relative to actual target position, with negative values indicating underestimation.

Percent error in VE conditions relative to that observed in the RE is also investigated; we call this transformation of the dependent variable *relative percent error* (RPE). This transformation is intended to better capture error in perceived distance, independent of task performance. Relative percent error was derived by first calculating mean RE percent error for each participant, task, distance, and device cell; this value was then subtracted from the percent error observed for each VR trial in the cell. Note that with regards to relative percent error, *device* is a factor of only the two VE condition levels (*vive*, *rift*).

The experiment progressed through measures in a random order; for each measure, the displays were presented in a random order; for each measure and display combination, a series of 9 distances (three repetitions of 2m, 3m and 4m), were presented in a random order. That is, each combination of measure, display, and distance was presented 3 times, for a total of 108 trials per participant. Measure was chosen as the factor at the top of the randomization hierarchy, as we were concerned that participants might forget how to perform a measure over time. Displayed target distances were $\pm 0.25m$ of the selected distance, so that the repetition might be less obvious; adjusted distances are called jittered target distances; unadjusted distances are called base target distances. After the experiment, a brief demographic survey and post-experiment interview was conducted.

3 Discussion and Analysis

Our results were analyzed using repeated measures ANOVA with a 5% significance level. When Mauchly's test indicated that the assumption of sphericity had been violated, Green-Geisser estimates of sphericity were used to correct degrees of freedom. Post-hoc analyses were performed using pairwise t-tests and a Holms-Bonferroni correction to achieve a 5% significance level. The re-

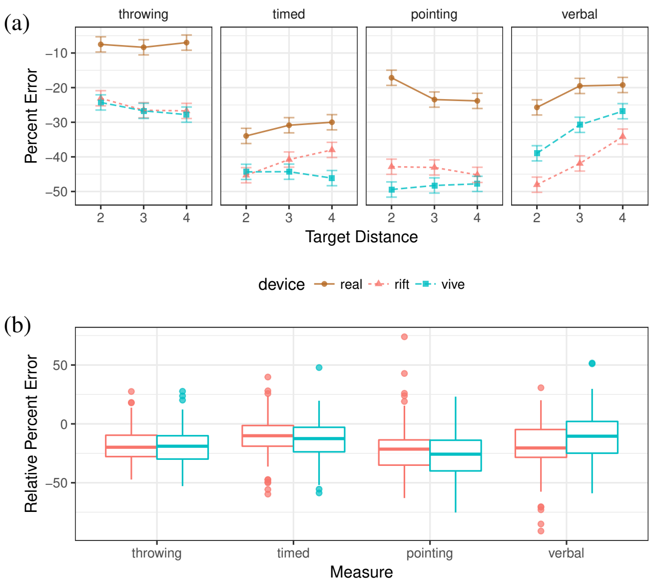

Figure 2: (a) Percent error (PE) of perceived distance. Target distances are base target distances. Error bars show 95% confidence intervals. (b) Relative percent error (RPE) by measure and device.

sults of these analyses are summarized in Table 2; descriptive statistics are summarized in Table 1.

We see clear underestimation in VE conditions. The effect of display on PE is seen to be significant ($F(2,32) = 49.06, p < 0.001$), with VE conditions showing 17% more mean underestimation than RE conditions ($p < 0.001$). This difference is clearly visible in Figure 2(a). PE for all devices and measures ranges from 22% to 47% mean underestimation. These results are in keeping with previous literature [17], suggesting that distance compression should be expected in contemporary consumer hardware.

Due to the large degree of underestimation in RE conditions, transforming PE into RPE leads to substantial changes. Underestimation is reduced to under 30% for all measures, and under 20% for all but one measure (pointing). Timed imagined walking is roughly tied for most underestimation as measured by PE ($M = -39.29, SD = 25.53$), but exhibits the least underestimation using RPE ($M = -11.53, SD = 16.41$). Distance estimates for all measures are more similar using RPE (compare Figures 2(a) and 2(b)), but still significantly different ($F(3,48) = 3.96, p = 0.013$). This suggests that reporting RPE as described here may facilitate comparing error in perceived distance across studies using the same measure, but that there are some additional, unaccounted for sources of variation between measures to consider in comparisons between measures. As RPE seems the more consistent measure of experienced compression across measures, the rest of our analysis considers only RPE and VE conditions.

The interaction effect between measure and distance ($F(6,96) = 2.86, p = 0.013$) suggests that all measures but timed behave differently across the range of distances we explore. Throwing seems to exhibit more underestimation at 4m than 2m ($p = 0.008$), pointing exhibits more underestimation at the nearest distance (2m) ($p = 0.002$), and verbal shows more underestimation at the outer edge (4m) (2m:$p = 0.008$, 3m:$p = 0.011$). This may be due to shifts in technique while performing measures (eg. changes in the physical demands of throwing to 2m versus 4m), differing mental models of the space influenced by cues presented by the real and virtual environments unique to this experiment, or per-measure effects that occur at the outer edges of a learned range. Further study using different environments and distances may help elucidate. Mean difference caused by these effects is relatively small, ranging from 4-6%.

Table 1: Summary of percent error in measured perceived distance per device and task.

Task	Display	N	Error (%) Mean	Error (%) SD	Relative Error (%) Mean	Relative Error (%) SD
throwing	real	153	-7.96	11.07	–	–
	rift	153	-24.60	14.13	-17.82	14.37
	vive	153	-25.12	14.75	-18.64	15.19
timed	real	153	-31.16	26.04	–	–
	rift	153	-41.31	24.69	-9.75	16.08
	vive	153	-45.04	19.35	-13.31	16.59
pointing	real	153	-22.01	9.64	–	–
	rift	153	-42.24	18.43	-22.23	20.67
	vive	153	-47.10	16.11	-27.02	18.50
verbal	real	153	-19.90	30.20	–	–
	rift	153	-39.64	31.68	-19.89	20.35
	vive	153	-29.54	32.67	-10.69	19.41

Table 2: ANOVA results of interest. Df_n and Df_d are the numerator and denominator degrees of freedom, F is the F-value, p is conditional probability of the F-test, η^2 is eta-squared, η_P^2 is partial eta-squared. All significant effects on RPE are shown; effects of measure and several interactions on PE are omitted for space.

Effect	On	Df_n	Df_d	F	p	Sig.	η^2	η_P^2
Display (D)	PE	2	32	49.06	$< .001$	***	.12	.75
Measure (M)	RPE	3	48	3.96	.013	*	.09	.20
M *x* D	RPE	3	48	6.53	.001	***	.03	.29
M *x* Distance	RPE	6	96	2.86	.013	*	.02	.15

Interaction effects between HMDs and measure ($F(3,48) = 6.53, p = 0.001$) are harder to interpret. The published specifications of both the Oculus Rift and HTC Vive are all but identical, yet the timed estimates performed using the Rift show less underestimation than those of all other measurement tasks by 7-12% ($p < 0.001$), while when using the Vive pointing shows more underestimation than all other measures by 9-17% ($p < 0.001$), and throwing differs from all measures but pointing by the more modest 5-8% (verbal:$p < 0.001$, throwing:$p = 0.002$). This may be due to the published difference in weight of 85g, though previous work suggests that HMD weight should not have an effect [22]; it be may due to unpublished specifications, such as the accommodative distance of the HMDs' lenses, as accommodative distance has been shown to influence distance perception in CAVEs [1]; it may be due to unknown deviations of our specific HMDs from manufacturer specifications. Issues in fit may also be an influence, as participants repeatedly put on and took off both HMDs, with only subjective confirmation of comfort and clarity of vision; one headset may have trended to some misalignment of participant's eyes to HMD lenses, or some gap between the HMD and face facilitating peripheral stimulus, as in [7]. None of these seems a satisfying explanation for the specific device-measure trends seen here, and further experiments may be needed to explore these effects.

One possible source of error, and a limitation inherent to our within-subjects design, are order effects. Previous work has shown order of presentation between real and virtual environments to influence distance estimations, in the specific contexts of single changes between environments, when VEs are a photorealistic match to the RE used [6,20,24,26]. Other learning or fatigue effects may also have been present. A simple linear regression on relative percent error over all VE trials yields a slope of 0.04, which would lead to a reduction of underestimation of 4.3% over a participant's 108 trials; RE trials yield a slope of 0.08, suggesting an 8.64% improvement. This suggests a slight overall learning effect twice as strong for RE as for VE, but no convergence between estimates in the two environments as might be suggested by environment presentation order effects [24, 26], and no strong overall bias suggesting interference from other order effects. This, along with our randomized presentation order and non-photorealistic VE, suggests that our results are not the product of an order effect. Further experiments using a counterbalanced or between-subjects design would further eliminate the possibility.

4 Conclusion

This paper evaluates four different measures of determining distance on two consumer grade HMD devices, within the tracked space provided by consumer-grade hardware. For all measures, significant differences between measurements made using virtual reality displays and the physical environment are observed, suggesting both that distance compression is induced by contemporary displays, and that consumer grade room-scale tracking can be used to measure distance compression. Evaluating measurements of underestimation made when viewing virtual environments relative to those made when viewing a real environment results in a 20% overall reduction of average underestimation and drastically transforms results under some measures, which suggests error in task performance may influence measured underestimation. Differences between measures remained statistically significant, suggesting that the use of relative percent error alone does not facilitate direct comparisons between measures; however, it may still be a valuable means of comparing results between studies using the same measure. Differences in underestimation between measures across distances, and between HMDs across measures, may merit further experiments isolating these effects for detailed exploration.

References

[1] G. Bruder, F. Argelaguet, A.-H. Olivier, and A. Lécuyer. Cave size matters: Effects of screen distance and parallax on distance estimation in large immersive display setups. *PRESENCE: Teleoperators and Virtual Environments*, (00), 2016.

[2] G. Bruder, F. A. Sanz, A.-H. Olivier, and A. Lécuyer. Distance estimation in large immersive projection systems, revisited. In *2015 IEEE Virtual Reality (VR)*, pages 27–32. IEEE, 2015.

[3] S. H. Creem-Regehr, J. K. Stefanucci, W. B. Thompson, N. Nash, and M. McCardell. Egocentric distance perception in the oculus rift (dk2). In *Proceedings of the ACM SIGGRAPH Symposium on Applied Perception*, pages 47–50. ACM, 2015.

[4] S. S. Fukusima, J. M. Loomis, and J. A. Da Silva. Visual perception of egocentric distance as assessed by triangulation. *Journal of Experimental Psychology: Human Perception and Performance*, 23(1):86, 1997.

[5] T. Y. Grechkin, T. D. Nguyen, J. M. Plumert, J. F. Cremer, and J. K. Kearney. How does presentation method and measurement protocol affect distance estimation in real and virtual environments? *ACM Transactions on Applied Perception (TAP)*, 7(4):26, 2010.

[6] V. Interrante, B. Ries, J. Lindquist, M. Kaeding, 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.

[7] J. A. Jones, J. E. Swan II, and M. Bolas. Peripheral stimulation and its effect on perceived spatial scale in virtual environments. *IEEE transactions on visualization and computer graphics*, 19(4):701–710, 2013.

[8] E. Klein, J. E. Swan, G. S. Schmidt, M. A. Livingston, and O. G. Staadt. Measurement protocols for medium-field distance perception in large-screen immersive displays. In *2009 IEEE Virtual Reality Conference*, pages 107–113. IEEE, 2009.

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

[10] B. Li, R. Zhang, and S. Kuhl. Minication affects action-based distance judgments in oculus rift hmds. In *Proceedings of the ACM Symposium on Applied Perception*, pages 91–94. ACM, 2014.

[11] B. Li, R. Zhang, A. Nordman, and S. A. Kuhl. The effects of minification and display field of view on distance judgments in real and hmd-based environments. In *Proceedings of the ACM SIGGRAPH Symposium on Applied Perception*, pages 55–58. ACM, 2015.

[12] B. J. Mohler, S. H. Creem-Regehr, and W. B. Thompson. The influence of feedback on egocentric distance judgments in real and virtual environments. In *Proceedings of the 3rd symposium on Applied perception in graphics and visualization*, pages 9–14. ACM, 2006.

[13] P. E. Napieralski, B. M. Altenhoff, J. W. Bertrand, L. O. Long, S. V. Babu, C. C. Pagano, J. Kern, and T. A. Davis. Near-field distance perception in real and virtual environments using both verbal and action responses. *ACM Transactions on Applied Perception (TAP)*, 8(3):18, 2011.

[14] A. Peer and K. Ponto. Perceptual space warping: Preliminary exploration. In *2016 IEEE Virtual Reality (VR)*, pages 261–262, March 2016.

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

[16] D. R. Proffitt, J. Stefanucci, T. Banton, and W. Epstein. The role of effort in perceiving distance. *Psychological Science*, 14(2):106–112, 2003.

[17] R. S. Renner, B. M. Velichkovsky, and J. R. Helmert. The perception of egocentric distances in virtual environments-a review. *ACM Computing Surveys (CSUR)*, 46(2):23, 2013.

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

[19] C. S. Sahm, S. H. Creem-Regehr, W. B. Thompson, and P. Willemsen. Throwing versus walking as indicators of distance perception in similar real and virtual environments. *ACM Transactions on Applied Perception (TAP)*, 2(1):35–45, Jan. 2005.

[20] F. Steinicke, G. Bruder, K. Hinrichs, M. Lappe, B. Ries, and V. Interrante. Transitional environments enhance distance perception in immersive virtual reality systems. In *Proceedings of the 6th Symposium on Applied Perception in Graphics and Visualization*, pages 19–26. ACM, 2009.

[21] 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*, 13(5):560–571, 2004.

[22] P. Willemsen, M. B. Colton, S. H. Creem-Regehr, and W. B. Thompson. The effects of head-mounted display mechanical properties and field of view on distance judgments in virtual environments. *ACM Transactions on Applied Perception (TAP)*, 6(2):8, 2009.

[23] P. Willemsen, A. A. Gooch, W. B. Thompson, and S. H. Creem-Regehr. Effects of stereo viewing conditions on distance perception in virtual environments. *Presence: Teleoperators and Virtual Environments*, 17(1):91–101, 2008.

[24] B. G. Witmer and W. J. Sadowski. Nonvisually guided locomotion to a previously viewed target in real and virtual environments. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 40(3):478–488, 1998.

[25] M. K. Young, G. B. Gaylor, S. M. Andrus, and B. Bodenheimer. A comparison of two cost-differentiated virtual reality systems for perception and action tasks. In *Proceedings of the ACM Symposium on Applied Perception*, pages 83–90. ACM, 2014.

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