Impact of Depth Perception on Virtual Reality Raycast Target Selection

ANONYMOUS AUTHOR(S)

We examine whether depth perception has a noticeable effect on raycast selection in VR. A controlled experiment examines the effect of 5 target “depths” defined as the distance from a set of planar targets to the user: 0.5m, 1.5m, 4.5m, 13.5m, and infinity (implemented as 10km). At all depths, tested angular widths and amplitudes are held constant in 1D and 2D versions of the 5-target Fitts' Law task. We also test two interaction methods: a typical laser pointer method and Janzen et al.'s aim-through-controller method. Results show that depth has no detectable effect on selection time, error rate, arm fatigue, and wrist fatigue; only eye strain was highest at the closest depth. Our results suggest that Fitts' Law and its variant models may not reflect this aspect of pointing behaviour, and we discuss the implications of our finding for VR researchers and designers.

Additional Key Words and Phrases: virtual reality, depth perception, raycast pointing

1 INTRODUCTION

Depth perception makes designing for VR different from other mediums. We define *depth* as the egocentric distance between the user's eyes and the target, and *depth perception* as an observer's ability to visually discern the distance of a given target [5]. We limit our discussion to VR based on head-mounted displays (HMDs) with a fixed focal length, like most commercially-available systems (*e.g.*, Meta Quest, Quest Pro, HTC VIVE, Sony PlayStation VR). Compared to typical real-world 2D displays, the stereoscopic and three-dimensional nature of VR adds depth perception as a potential factor affecting interaction performance. In particular, raycasting is a popular method of VR target selection, so understanding how depth perception affects raycasting performance is important, with practical implications for both researchers and industry practitioners.

Design recommendations from Meta [1] and Microsoft [14] recommend placing targets intended for prolonged use around 1-2 metres away from the user to match the virtual (from stereoscopic rendering) and physical (from the headset's optic focal point) depths of targets. A likely reason is to reduce the user's vergence-accommodation conflict, which can cause unpleasantness and visual fatigue [9]. However, this brings about several practical questions. How much does the mismatch of virtual-physical depth matter? Could a designer place content extremely far away (*e.g.*, to create an ambient menu or scoreboard in the skybox) without significantly sacrificing interaction performance or comfort? These questions are difficult to answer from industry guidelines, warranting a more nuanced exploration.

Janzen et al. [11], who play an important role in motivating our study, modelled raycast selection performance in relation to target depth. They found that, given an increase in depth, the relative influence of changes in *target width* on selection performance increased compared to the influence from the changes in *movement amplitude* (distance between targets). However, the authors controlled the absolute linear dimensions in space, not angular dimensions relative to the user's viewpoint. This means that the targets became visually smaller and closer together as depth increased. As a result, it is hard to conclusively deduce the isolated effect of depth separate from the changes in apparent angular target width and amplitude. Kopper et al. [12] suggests angular dimensions, rather than their linear counterparts, are more accurate factors of interaction performance. While their claim was challenged by Janzen et al. [11], we revisit this idea to see if controlling the angular dimensions of targets influences the effect of depth on raycast interaction. Our primary research question is: *what is the effect of depth on raycast selection if visual target width and amplitude are held constant?*

. Manuscript submitted to ACM

To explore this, we ran a controlled lab experiment, where participants performed a standard 1D horizontal and 2D circular Fitts' Law selection task [10] with varying depths. The set of tested target widths and amplitudes are held constant in the angular dimension (measured in degrees) across all conditions. The experiment also compares whether the effect of depth on selection changes based on the raycast techniques. The first method replicates Janzen et al.'s aim-through-controller pointing method in which the pointing ray originates from the user's head and goes through the controller, and the second method uses a more typical VR pointing method where the ray originates from the controller itself. We found that the target depth did not affect selection performance across all techniques in all measures (selection time, error rate, arm fatigue, and wrist fatigue) except for eye strain, which was higher for the closest depth $D = 0.5m$. We compare our findings with previous work and deduce that depth only affects selection in that it *scales* the target width and amplitude. Fitts' Law and its variant models do not reflect the influence of this depth-dependent target scaling on selection, which suggests that future work should explore this discrepancy and how current models could be modified to reflect the effect of scaling. From a design implication perspective, our results show that the choice of target depths can be chosen according to design needs without affecting selection performance or comfort as long as targets are not too close to the user's eyes ($> 0.5m$).

We make two main contributions: (1) an empirical investigation demonstrating that target depth alone has no detectable impact on raycast target selection when the angular dimensions are held constant; and (2) design implications for VR user interface designs that are enabled by exploiting a large range of target depths, for different aesthetic or immersive goals.

2 RELATED WORK

We discuss how current standard practices of VR designs take depth into consideration. We then present previous work that explores the effect of depth perception on VR raycasting and identify gaps in existing research.

2.1 Standard Practice for Target Depth in VR

There exist industry recommendations for target depths in VR. Meta [1] recommends 1m and Microsoft [14] 2m to reduce the chance of nausea from vergence-accommodation conflict. *Vergence-accommodation conflict* is caused by the disparity between accommodative demand (focusing of images by altering eye lens shape) and vergence demand (the extent to which the eyes rotate inwards to allow the lines of sight to intersect at the desired depth plane) [9]. Typical VR headsets use a fixed focal plane, meaning accommodation does not change, while vergence varies across the virtual depth of the target. A mismatch can cause unpleasantness and eye strain [9]. However, the exact extent to which this negatively affects the VR experience and whether it matters for interactions with relatively short durations (as is the case with target selection for interfaces like menus) is unclear. Moreover, could the conflict have no effect on objective selection performance (*i.e.*, time and error rate) despite some effect on the subjective strain and comfort? Additionally, even if such a conflict does affect selection, vergence ceases to change for targets farther than 10m away [13]. This suggests that it is interesting to study how depth affects selection in extremely far distances (such as 10m vs. infinity, *i.e.* the environment's skybox).

Standard guidelines also exist for academic studies. Bergström et al. [4] provide a set of recommendations for VR studies based on a review of 20 years of work. They found that the farthest target depth used in a study was 10m from Yu et al. [20]. This further illustrates that exploring a much greater depth could be interesting. They also suggest that angular widths and amplitudes of the targets should be reported. Presumably, the goal is to make results comparable across different experiments. But this makes an implicit assumption that if targets have the same visual size and amplitude, depth itself does not introduce a confound. However, no prior work contributes empirical evidence that such a premise is true. Some previous work shows that depth may affect selection performance in other mediums like large stereoscopic displays [11], as well as other interactions like direct touch [3, 16] and size manipulation [2].

2.2 Depth Perception and VR Raycasting

Janzen et al.'s study [11] on raycast selection used a planar large stereoscopic display, and found that depth correlates with $k = \frac{b_2}{b_1}$, a parameter introduced by Shoemaker et al. [15] representing the ratio between the influence of target width and amplitude on selection time in Welford's model of selection performance [17]:

$$\text{Selection Time}(T) = a + b_1 \log_2(A) - b_2 \log_2(W), \quad (1)$$

where a, b_1, b_2 are model parameters, and W (ω in angular dimension) and A (α in angular dimension) are target width and amplitude, respectively. More intuitively, Janzen et al.'s finding shows that as depth increases, changes in W have more influence on selection time than changes in A. They also use other similar models (all variants of Fitts' Law [8]), but we use the one above for ease of reasoning. A similar correlation with k was observed by Shoemaker et al. [15] between *gain* (ratio between the movement of the input device and the display pointer [7]) and selection time on 2D displays. Interestingly, no theory-based explanation exists for why this effect on k exists. However, as noted in the introduction, Janzen et al.'s A and W conditions were controlled linearly (in centimetres). As the targets move further, they appear smaller and closer together (*i.e.*, lower α and ω). Therefore, the influence of depth could be merely limited to how it scales α and ω. This reasoning for the effect of depth on selection performance also extends to gain. An increase in gain makes the pointer cursor move faster, which requires the user to move the pointing device slower and more precisely. We reason that this effect of gain is equivalent to shrinking the targets and reducing the movement amplitude while keeping the gain constant. Hence we propose a question: If we control the apparent target width and amplitude (*e.g.*, $\omega = 3°, \alpha = 15°$) to be the same across different depths, would depth have any effect on selection performance? In other words, does depth have any meaningful impact on selection other than that, when we fix the linear A and W, it visually scales the apparent target width and amplitude?

Let us examine potential factors of depth perception that may affect selection other than the visual scaling of targets. Regarding monocular perception (how the targets look in each eye), targets with constant angular width appear nearly identical at all depths. Accommodation, an oculomotor cue stemming from the changes in the eye's focus, cannot act as a depth cue due to the fixed focal length used in a VR headset. A notable binocular depth cue is vergence, which describes how the eyes rotate and converge to focus on a target. The effective range of vergence is up to 10m [13]. In other words, depth perception is limited past this distance, with only monocular environmental cues available to convey depth information, such as relative size and perspective from parallel lines [13]. So, if there is indeed an effect of depth perception on raycast selection, we may see differing extents of such an effect between close and far depth ranges. However, these are simply cues for depth perception. Regardless of whether the user can distinguish the depth of the targets or not, we are not aware of any intuitive or information-theoretic reason why selection performance would differ by depth on targets with constant apparent width and amplitude. One potential cause of eye strain or discomfort is the vergence-accommodation conflict, but whether that will noticeably harm selection performance is still an open question.

Interestingly, Janzen et al. reports that using angular dimensions α and ω did not improve the fitness of the model compared to the linear dimension counterparts [11], contradicting the claim by Kopper et al. that angular dimensions are more accurate factors of interaction performance [12]. Intuitively, if the targets appear approximately the same

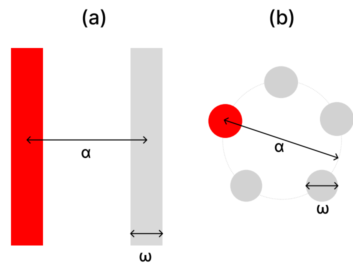

Fig. 1. Fitts' Law tasks: (a) 1D (b) 2D

from the user's perspective, it seems reasonable that their depth should have little or no discernible impact on selection. We aim to resolve these conflicting views by means of an empirical lab experiment.

3 EXPERIMENT

In this in-lab experiment, we investigated raycast selection performance in VR with varying target depths, with a constraint that the set of angular (rather than linear) target widths and amplitudes is kept constant. We test 1D and 2D Fitts' Law tasks with two pointing techniques: a typical laser-pointer method and the aim-through-controller method used by Janzen et al.. If depth results in a deviation in pointing motor behaviour with the typical technique (*e.g.*, closer targets act more like a direct touch interaction than raycasting), then we expect to observe a difference in our measures between the two methods (*i.e.*, an interaction between the pointing techniques and depth). We analyze selection time, error rate, and subjective fatigue ratings.

3.1 Participants

We recruited 20 right-handed participants (ages 17-28, 18 male, 2 female), which follows recommendations for VR studies by Bergström et al. [4]. Recruiting used word-of-mouth and platforms like Reddit, Facebook Groups, and Discord. Most participants were either novice or intermediate VR users: 3 never used VR, 12 use it a few times a year, 2 a few times a month, 1 a few times a week, and 2 use it daily. Participants reported an average interpupillary distance (IPD) of 63.7mm ($\sigma = 3.6mm$). Each participant received a $15 gift card as remuneration.

3.2 Apparatus

The experiment software is built with Unity version 2020.1.6f1 and the SteamVR framework for tracking. Raycast pointing and target selection use the Vive Input Utility [19], with tracking data captured at least 72 times per second. We used the Meta Quest Pro headset along with a high-performance computer with an RTX 3080 GPU.

3.3 Task

Participants performed a 1D reciprocal 2-target task and a 2D circular 5-target ISO 9241-411 selection task (Figure 1) [10]. The current active target is highlighted in red, with all others grey. Participants select the active target using their index finger trigger button. Upon selection, the next target is highlighted, which is the opposite bar target in the 1D task and across the ring clockwise to the previous target in the 2D task. Following Kopper et al. [12], participants are not required to correctly select the target to move on to the next trial. Upon each selection, a timestamp and correctness flag are recorded. The system emits a ding to indicate success and a beep to indicate an incorrect selection. Each selection counts as one *trial*, and each *repitition* (or *rep*) consists of 5 trials.

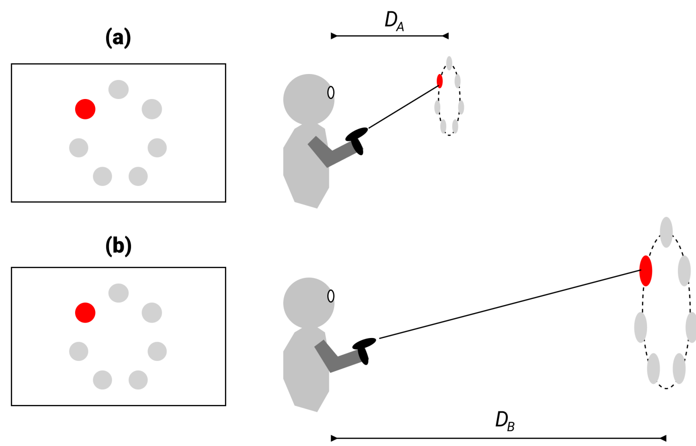

Fig. 2. The task varies target depth (distance from user to target plane D) with same apparent target width and same movement amplitude: (a) user view with shorter depth D_A; (b) targets appear the same to user regardless of greater depth D_B.

The center of the target plane is placed at the same height as the HMD at the beginning of the experiment, and rotated to face directly at the participant. The depth represents the distance between the HMD and the plane center, and it does not affect the apparent target width and movement amplitude (Figure 2). Mathematically, $W \propto D$ such that ω is kept constant (same for A and α). We control for user head movement by displaying a warning message if they move more than 10cm from their starting position.

The software renders a floor grid and a mountain range skybox for parallax to ensure that the users have environmental cues for depth perception. Users are seated and perceive themselves as sitting above the floor grid. Since the floor would obstruct the user's view of lower targets at large depths with large target sizes, we use a stencil buffer to render targets without the grid overlapping them.

3.4 Techniques

Participants used two raycast techniques: a more typical laser-pointer method (TYPICAL) and an aim-through-controller method used by Janzen et al. (JANZEN) [11]. The latter serves two purposes: (1) it replicates Janzen et al.'s experiment in HMD VR, and (2) it causes a controlled condition where α and ω are the same from the perspective of the user *and* of the controller.

In 2D tasks, with TYPICAL, participants point using a white ray emanating from the controller. A spherical raycast cursor is placed at the end of the ray with an apparent angular width of 0.2°. With JANZEN, the cursor is placed on the target plane along the line of sight between the centre of the eyes and the top of the controller. Here, only a cross-shaped cursor is shown without the controller or a ray (since it emanates from the user's eye).

In 1D tasks, everything remains the same except no ray is shown since it is an infinite plane, and the cursor is a thin vertical line with an apparent angular width of 0.1°.

3.5 Procedure

At the start of a new task or technique, the experimenter explained how to perform the mechanism of the raycast technique and which target to select. The participant was instructed to aim for a ~5% error rate, and that the first 20 trials (4 reps) were for practice. Between each condition, the system paused for a rest break and showed the participant their error rate for the previous trials. After they completed all 5 depths, participants rated their eye strain, arm fatigue, and wrist fatigue on a 7-point numerical scale. At the end of the experiment, there was a short interview for feedback and comments. Each experiment session took around 1 hour.

3.6 Design

This is a within-subjects design with target depth D as the main independent variable with 5 levels (0.5m, 1.5m, 4.5m, 13.5m, 10000m representing infinity), with TASK (1D, 2D) and TECHNIQUE (TYPICAL, JANZEN) as secondary factors. To test a variety of target configurations, there are 2 target width (ω) conditions (1° and 3°) and 2 movement amplitude (α) conditions (15.7° and 38.0°). Each combination of the primary and secondary independent variables was repeated for 4 REPS.

The order of TASK and D are counterbalanced using a 4×4 Latin square, and for each combination, the participant completes a randomly-ordered set of D combinations. Within each D condition, the order of ω and α are randomized.

Dependent measures are computed from logs. *Selection Time* is defined as the duration from the moment the index trigger is pressed until the next selection event. Selections outside the correct target boundary are also counted, as in previous experiment designs [3, 12, 15]. *Error Rate* is the ratio between the number of unsuccessful selections and the total number of target selections. Subjective feedback used 7-point numeric ratings for *Eye Strain*, *Arm Fatigue*, and *Wrist Fatigue*.

We conducted a power analysis to ensure that we have enough data to detect sufficiently significant differences. As recommended for experimental psychology studies [6], we aim for a power of 0.8 with a significance level of 0.05 even when the effect size is small ($d = .1$). With 5 D $\times$ 2 TASK $\times$ 2 TECHNIQUE = 20 conditions, the analysis shows the necessary number of observations is $n = 104$. For just depth (5 conditions), the number increases to $n = 240$. Since we have 20 participants $\times$ 2 ω $\times$ 2 α $\times$ 4 REPS $\times$ 5 trials = 1600 observations for each condition, the design has sufficient power.

4 RESULTS

For each combination of D, TASK, and TECHNIQUE, trial times more than 3 standard deviations from the mean selection time were excluded as outliers. In total, 50 trials (1.08%) were removed.

In the analysis to follow, a D $\times$ TASK $\times$ TECHNIQUE ANOVA with Holm-Bonferroni corrected post hoc pairwise t-tests was used, unless noted otherwise. We only report effects involving the main factor D since those are what we are most interested in. When the assumption of sphericity was violated, degrees of freedom were corrected using Greenhouse-Geisser ($\epsilon < 0.75$) or Huynh-Feldt ($\epsilon \geq 0.75$). Shapiro-Wilk tests showed that residuals for only about half the conditions exhibited normality. As a result, we run repeated-measures ANOVAs using Box-Cox transformed *Time* and aligned-rank transformed [18] measures for all others. For each measure, trials were aggregated by participant and factors being analysed. We also aggregate ω and α for analysis. Time data is aggregated by REP using the median to account for skewed distribution.

4.1 Learning Effect

We are interested in practised performance, so we examine if earlier repetitions took longer and should be removed. There is a significant main effect for REP on *Selection Time* ($F_{3,57} = 44.04$, $p < .0001$, $\eta^2_G = .03$), but no interaction effects

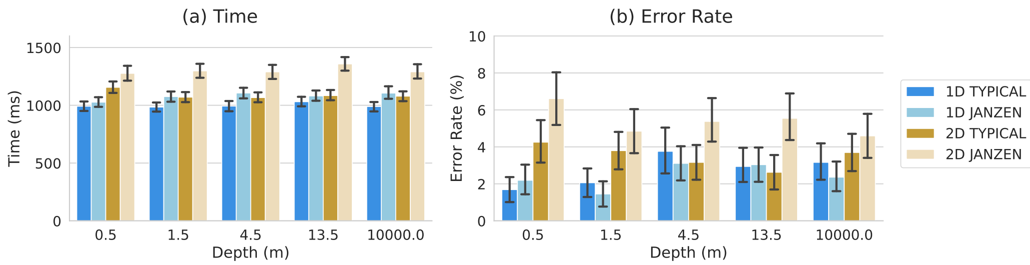

Fig. 3. (a) *Selection Time* (b) *Error Rate* by target depth D for each TASK ×SURFACE (error bars in all graphs are 95% confidence)

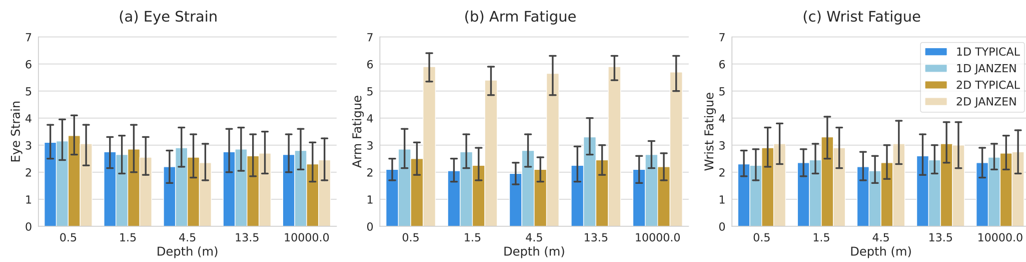

Fig. 4. (a) *Eye Strain* (b) *Arm Fatigue* (c) *Wrist Fatigue* by target depth D for each TASK ×SURFACE

involving BLOCK. Post hoc tests found repetition 1 significantly slower than other repetitions ($p < .001$). In subsequent analysis, only repetitions 2, 3 and 4 are included to be more representative of practised performance.

4.2 Time and Error Rate

Depth did not affect selection time or error rate (Figure 3). There was no main effect of D on *Selection Time* ($F_{4,76} = 1.84$, $p = .13$) or *Error Rate* ($F_{4,4700} = 1.02$, $p = .40$), and no TASK or TECHNIQUE interaction effects.

4.3 Subjective Ratings

The closest depth (0.5m) caused greater eye strain than farther depths (Figure 4a). There was a main effect of D on *Eye Strain* ($F_{4,361} = 7.20$, $p < .0001$), but no interactions involving D. Post hoc tests show D = 0.5M ($\bar{x} = 3.2, \sigma = 1.7$) was rated as higher eye strain than D = 1.5M ($\bar{x} = 2.7, \sigma = 1.7$), 4.5M ($\bar{x} = 2.5, \sigma = 1.7$), 13.5M ($\bar{x} = 2.7, \sigma = 1.9$), and INF ($\bar{x} = 2.6, \sigma = 1.7$) (all $p < .005$).

For arm or wrist fatigue, there were no main effects or interactions involving D (Figure 4b,c).

5 DISCUSSION

We discuss the experiment results in light of insights from previous research, outline how HCI researchers and designers can use our findings, acknowledge the limitations of our study, and describe potential future work.

Depth Does Not Directly Affect Target Selection. Our results found target depth does not significantly affect selection time or errors when targets are scaled to have the same apparent width and amplitude. We expected to observe some difference in the TYPICAL technique, but we did not. This suggests that any changes in pointing motor behaviour due to target depth do not affect selection performance. This contrasts with Janzen et al. [11], who report that there exists an effect of depth in a non-VR one-dimensional target selection task. We explain this mathematically to reveal an interesting insight from this apparent disagreement.

Let s be a scalar that proportionally scales the angular movement amplitude α and target angular width ω, such that $\alpha' = s\alpha$ and $\omega' = s\omega$. Note that s can be calculated for any depth D such that the perceived visual angle of α' and ω' remain constant from the user's viewpoint. When this relationship between s and D exists, then the target selection time T is independent of D.

Recall our discussion in Section 2.2 on Janzen et al. [11]'s suggestion that $k \propto D$, where $k = \frac{b_2}{b_1}$ in the Welford model [17]: $T = a + b_1 \log_2(A) - b_2 \log_2(W)$. Their result seems to contradict ours; depth should not affect the behaviour of selection performance. The reason for this conflict is likely because Janzen et al. use linear target width W and amplitude A, which entails that $s \propto \frac{1}{D}$. In other words, the *apparent size* of the target widths (ω) and amplitudes (α) that are further away appear smaller to the user. So, if $k \propto D$ and $s \propto \frac{1}{D}$, then it should follow that

$$\therefore k \propto \frac{1}{s}. \tag{2}$$

In other words, in Janzen et al.'s study, depth (and, by extension, gain in Shoemaker et al.'s study [15]) had its effect on k only because it scales α and ω. Therefore, **depth has no inherent effect on selection performance other than that it typically scales the target width and amplitude.**

Incorporating s and the angular α and ω into the Welford model reveals an interesting conflict:

$$\begin{aligned} T &= a + b_1 \log_2(s\alpha) - b_2 \log_2(s\omega) \\ &= (a + b_1 \log_2(s) - b_2 \log(s)) + b_1 \log_2(\alpha) - b_2 \log_2(\omega). \end{aligned} \tag{3}$$

This produces a new term $a' = a + b_1 \log_2(s) - b_2 \log(s)$ that is a function of s. We can see that s does not influence b_1 or b_2; hence, the model implies that k is independent of s. Shoemaker et al. [15] also noted this characteristic of the model. Our finding that $k \propto \frac{1}{s}$ challenges this result implied by the model and opens a new research question to be explored. We can observe a similar behaviour with other variant models of Fitts' Law: the coefficients of the logarithm terms for α (or A) and ω (or W) are unaffected by s.

Furthermore, our finding contributes evidence supporting the idea that the original Fitts' notion of index-of-difficulty ($ID = \log_2(\frac{A}{W})$ or $\log_2(\frac{\alpha}{\omega})$) [8] is an inaccurate model for pointing performance [11, 15, 17]. Any scaling s of α and ω is cancelled out, which contradicts our results that the absolute apparent target scale (separate from the relative ratio between α and ω) matters in selection.

Only Closest Depth Leads to Higher Eye Strain. Eye strain is greater at the closest depth $D = 0.5m$, while arm and wrist fatigue was unaffected. This result makes sense because converging the eyes on a very close target involves the vergence-accommodation conflict, which is known to cause eye fatigue [9]. Therefore, we recommend avoiding very close targets of depths of 0.5m or less. As for targets farther away, depths from 1.5m onward had no adverse effect despite the expected vergence-accommodation conflict [1, 14]. This entails that designers can be creative and use targets at these depths without compromising performance or comfort. Interestingly, participants' verbal feedback was highly variable: while the majority said that the eye strain and fatigue were mostly the same across depths, some said that fatigue got better with targets farther away, while others said they felt better with targets closer to the "medium" distance of 1.5m.

5.1 Design Implications

Our findings can inform experiment design, interaction techniques and interface designs.

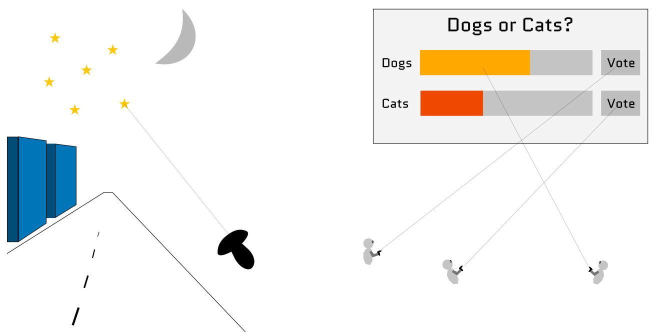

Fig. 5. Example VR interactions enabled by placing targets at very far depths: (a) raycast selection of targets placed on the skybox itself; (b) shared billboard voting menu for consistent visualization and input for many VR players.

5.1.1 *Researchers.*

When designing raycast selection experiments in VR, our findings suggest that target depths will not act as a confound as long as they are reasonably separated from the user's view, further away than arm's reach (*i.e.* depths further than our $D = 0.5m$ condition). While target depth should be reported for replicability, there is no evidence that different depths (especially medium versus far) affect a selection performance evaluation as long as the target apparent sizes and movement amplitudes are kept constant throughout. We recommend that target widths and amplitudes be reported, controlled, and analyzed using angular measures. This expands the recommendations for VR experiments from Bergström et al. [4].

5.1.2 *Application Designers.*

Our findings indicate that objects and menus in VR for raycast selections can be placed very far away or as close as necessary to nearly arm's reach to accommodate raycast interaction design needs without sacrificing selection performance or workload. This is not to say that target depths should not be considered, but rather that the desired design outcome can dictate much of these measures with greater room for creativity. For example, in VR games, one can place raycast targets on the skybox to enable various playful interactions, such as making the moon and stars part of an in-game menu system (Figure 5a). A very large community billboard can be created that everyone can publicly view and interact with as a single shared experience (Figure 5b).

5.2 Limitations

Each depth condition lasted approximately 4-5 minutes, with a break every 4 reps (20 trials). Oculus and Microsoft recommend 1-2m as the target depth intended for use in "extended periods of time" [1, 14]. It is possible that an effect of depth on subjective fatigue ratings might appear with a much longer task period. However, we believe that our task period length is a fair representation of real VR usage, such as the maximum time a player might focus on a planar set of targets (for an in-game menu). Additionally, longer task lengths would have resulted in excessive time requirements for the participants, making recruiting and logistics challenging, as well as risking unncessary fatigue.

5.3 Future Work

Future work could explore VR user interface designs enabled by extremely large target depths and subjective preferences of such a target placement. We briefly describe two examples in the previous section, but whether those are preferred over the current more common designs in VR applications is unknown.

Our results also warrant more investigation into the discrepancy between our findings (k and s are correlated) and the Welford model (k and s are independent). Future studies could, for instance, control the same A-to-W ratio, scale their apparent size, and model the behaviour of k. Moreover, how the Welford model could be modified to reflect the effect of s on k is an open question for future work to explore.

6 CONCLUSION

Our investigation suggests that depth has very little to no inherent effect on selection time or error rate in VR raycast target selection tasks, nor any pronounced effect on subjective eye strain and physical fatigue. Any effect demonstrated in previous work is likely due to how a change in depth scales the targets accordingly when target width and amplitude are linearly controlled. We hope that our findings can inform future VR experiment design and lead to novel and innovative VR interface design, perhaps even unlocking new ways for people to interact in VR without sacrificing performance or comfort.

REFERENCES

[1] 2021. Vision. https://developer.oculus.com/learn/bp-vision/

[2] Rahul Arora, Jiannan Li, Gongyi Shi, and Karan Singh. 2021. Thinking Outside the Lab: VR Size & Depth Perception in the Wild. *arXiv e-prints*, Article arXiv:2105.00584 (May 2021), arXiv:2105.00584 pages. arXiv:2105.00584 [cs.HC]

[3] Mayra Donaji Barrera Machuca and Wolfgang Stuerzlinger. 2019. *The Effect of Stereo Display Deficiencies on Virtual Hand Pointing*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3290605.3300437

[4] Joanna Bergström, Tor-Salve Dalsgaard, Jason Alexander, and Kasper Hornbæk. 2021. How to Evaluate Object Selection and Manipulation in VR? Guidelines from 20 Years of Studies. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 533, 20 pages. https://doi.org/10.1145/3411764.3445193

[5] Myron L Braunstein. 2014. *Depth perception through motion*. Academic Press.

[6] Marc Brysbaert and Michaël Stevens. 2018. Power Analysis and Effect Size in Mixed Effects Models: A Tutorial. *Journal of Cognition* (Jan 2018). https://doi.org/10.5334/joc.10

[7] Géry Casiez, Daniel Vogel, Ravin Balakrishnan, and Andy Cockburn. 2008. The impact of control-display gain on user performance in pointing tasks. *Human–computer interaction* 23, 3 (2008), 215–250.

[8] Paul M Fitts. 1954. The information capacity of the human motor system in controlling the amplitude of movement. *Journal of experimental psychology* 47, 6 (1954), 381.

[9] David M. Hoffman, Ahna R. Girshick, Kurt Akeley, and Martin S. Banks. 2008. Vergence-accommodation conflicts hinder visual performance and cause visual fatigue. *Journal of Vision* 8 (3 2008), 33–33. Issue 3. https://doi.org/10.1167/8.3.33

[10] ISO 9241-411:2012 2012. *Ergonomics of human-system interaction — Part 411: Evaluation methods for the design of physical input devices*. Standard. International Organization for Standardization, Geneva, CH.

[11] Izabelle Janzen, Vasanth Kumar Rajendran, and Kellogg S. Booth. 2016. Modeling the Impact of Depth on Pointing Performance. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, San Jose, CA, USA, May 7-12, 2016*, Jofish Kaye, Allison Druin, Cliff Lampe, Dan Morris, and Juan Pablo Hourcade (Eds.). ACM, 188–199. https://doi.org/10.1145/2858036.2858244

[12] Regis Kopper, Doug A Bowman, Mara G Silva, and Ryan P McMahan. 2010. A human motor behavior model for distal pointing tasks. *International journal of human-computer studies* 68, 10 (2010), 603–615.

[13] Takanori Okoshi. 2012. *Three-dimensional imaging techniques*. Elsevier.

[14] Erick J Paul, Formicola Stefano, and Coulter David. 2020. Comfort - Mixed Reality. https://docs.microsoft.com/en-us/windows/mixed-reality/design/comfort#guidance-for-holographic-devices

[15] Garth Shoemaker, Takayuki Tsukitani, Yoshifumi Kitamura, and Kellogg S. Booth. 2012. Two-Part Models Capture the Impact of Gain on Pointing Performance. *ACM Trans. Comput.-Hum. Interact.* 19, 4, Article 28 (Dec. 2012), 34 pages. https://doi.org/10.1145/2395131.2395135

[16] Robert J. Teather and Wolfgang Stuerzlinger. 2013. *Pointing at 3d Target Projections with One-Eyed and Stereo Cursors*. Association for Computing Machinery, New York, NY, USA, 159–168. https://doi-org.proxy.lib.uwaterloo.ca/10.1145/2470654.2470677

[17] Alan Traviss Welford. 1968. Fundamentals of skill. (1968).

[18] Jacob O. Wobbrock, Leah Findlater, Darren Gergle, and James J. Higgins. 2011. *The Aligned Rank Transform for Nonparametric Factorial Analyses Using Only Anova Procedures*. Association for Computing Machinery, New York, NY, USA, 143–146. https://doi.org/10.1145/1978942.1978963

[19] Lawrence Wong. 2021. ViveSoftware/ViveInputUtility-Unity. https://github.com/ViveSoftware/ViveInputUtility-Unity

[20] Chun Yu, Yizheng Gu, Zhican Yang, Xin Yi, Hengliang Luo, and Yuanchun Shi. 2017. *Tap, Dwell or Gesture? Exploring Head-Based Text Entry Techniques for HMDs*. Association for Computing Machinery, New York, NY, USA, 4479–4488. https://doi.org/10.1145/3025453.3025964