Effect of Stereo Deficiencies on Virtual Distal Pointing

Anil Ufuk Batmaz
ufuk.batmaz@concordia.ca
Concordia University
Montreal, Canada

Moaaz Hudhud Mughrabi
moaaz.mazen@stu.khas.edu.tr
Kadir Has University
Istanbul, Turkey

Mayra Donaji Barrera Machuca
mbarrera@dal.ca
Dalhousie University
Halifax, Canada

Wolfgang Stuerzlinger
w.s@sfu.ca
Simon Fraser University
Vancouver, Canada

ABSTRACT

Previous work has shown that the mismatch between disparity and optical focus cues, i.e., the vergence and accommodation conflict (VAC), affects virtual hand selection in immersive systems. To investigate if the VAC also affects distal pointing with ray casting, we ran a user study with an ISO 9241:411 multidirectional selection task where participants selected 3D targets with three different VAC conditions, no VAC, i.e., targets placed roughly at 75 cm, which matches the focal plane of the VR headset, constant VAC, i.e., at 400 cm from the user, and varying VAC, where the depth distance of targets changed between 75 cm and 400 cm. According to our results, the varying VAC condition requires the most time and decreases the throughput performance of the participants. It also takes longer for users to select targets in the constant VAC condition than without the VAC. Our results show that in distal pointing placing objects at different depth planes has detrimental effect on the user performance.

CCS CONCEPTS

• **Human-centered computing → Pointing**.

KEYWORDS

Virtual Reality, Vergence-Accommodation Conflict, Stereo Deficiencies, Distal Pointing, Ray Casting, Selection

ACM Reference Format:
Anil Ufuk Batmaz, Moaaz Hudhud Mughrabi, Mayra Donaji Barrera Machuca, and Wolfgang Stuerzlinger. 2022. Effect of Stereo Deficiencies on Virtual Distal Pointing. In *28th ACM Symposium on Virtual Reality Software and Technology (VRST '22), November 29-December 1, 2022, Tsukuba, Japan.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3562939.3565621

1 INTRODUCTION

Thanks to recent innovations in virtual reality (VR) technologies, immersive systems have become more prevalent, affordable, and accessible, and are being used by a wider audience. For example, sectors like VR games, are expected to reach global revenues of 2.4 billion U.S. dollars by 2024 [10]. Current VR head-mounted displays (HMDs) are self-contained, wireless, light, and offer high visual fidelity, e.g., the Quest 2. Some HMDs even afford 4k resolution with low latency in rendering and tracking, wide field-of-views (FOVs), and an adjustable inter-pupillary distance (IPD), i.e., the distance between the center of the pupils of the eyes. Examples of such VR HMDs include the XR-3 Varjo HMD [46] and the Pimax 4k [39]. Despite all these advances, previous work has found that users still cannot interact as fast as with the mouse and touchscreens when selecting 3D objects [12], as measured by longer selection times and lower throughput. There are multiple potential explanations for this lower performance, including jitter [38, 43], bio-mechanical limitations [30, 41], and the way current stereo displays render content [3, 4].

In this paper, we focus on the latter explanation, i.e., the effect of conflicting depth cues caused by how current stereo display systems show content. To display 3D content, stereo displays show two different images to the users' eyes from viewpoints that correspond to the two eye positions in a human head. Each image is displayed/projected and focused at a fixed plane by the headset, typically through a 2D screen in VR HMDs. When displaying 3D content that is not at the same depth as said fixed plane, a user's eye is thus exposed to a mismatch between focusing on the display plane (accommodation) and rotating the eyes to see the object at its correct visual depth (vergence). This vergence-accommodation conflict (VAC) does not occur for targets in the real world, nor does it happen for targets that are on the focal plane. In other words, we posit that the way VR HMDs render the 3D content affects interaction with that content.

Here, we aim to quantify the effect of the VAC on 3D selection of distal targets using ray casting-based selection in current stereo display systems, using a similar virtual scene as in previous work [5], but with a different experimental design and task. Knowing about the effect of the VAC on ray-based interaction is important because ray casting is a popular interaction technique in current VR systems, e.g., to interact with menus and/or far away objects. Also, far objects appear smaller due to perspective, which makes them potentially harder to select. This can potentially increase any challenges introduced by the presence of VAC. By comparing selection of targets with ray casting at different visual depths, we can identify if the presence of VAC affects the interaction. Quantifying the effect of the VAC on user performance with further away targets, *out of*

VRST '22, November 29-December 1, 2022, Tsukuba, Japan

ACM ISBN 978-1-4503-9889-3/22/11…$15.00
https://doi.org/10.1145/3562939.3565621

arm's reach, i.e., more than ≈70 cm from the user, is important as in many consumer-level VR/AR applications users interact with objects that are further away. Also, current VR headsets often have a focal plane that is beyond arm's reach, i.e., further than 70 cm.

Our work on ray casting selection techniques thus extends previous work on depth perception that identified that stereo display deficiencies affect virtual hand interaction [3, 5], i.e., for objects within arm's reach. It also extends work that investigates the VAC [4, 21, 29, 31].

2 PREVIOUS WORK

2.1 Ray based selection techniques

In this paper, we focus on distal selection of 3D targets using the ray casting technique, which was introduced by Bowman et al. [8] in 1997. Ray casting allows the distal, i.e., remote selection of 3D object by using a ray that extends from the controller. When this ray intersects with an object in space and the user presses a trigger, the object is selected. This technique enables easy remote selection of targets and affords accurate selection at shorter distances. The main limitation of ray casting is the effect of unintentional hand tremor and/or tracker orientation variations, i.e., jitter, that affect the precision of the selection [38, 38, 43]. This issue is magnified when selecting far away targets, as small rotational displacements then lead to large-scale absolute translations [9]. Another problem that affects ray casting is that targets that are very close can require large angular movements [30].

To address these issues, research has proposed multiple solutions, including using disambiguation mechanisms like a depth-aware bubble cursor [19] and the RayCursor [1]. Another solution is adding additional degrees of freedom (DOF) to the ray casting technique. Examples of such interactions that add a single DOF, include ray casting with reeling [8] and Sun et al.'s work [42]. Other work added multiple DOF, e.g., Plane-casting [27] and INSPECT [28]. Although all these techniques improve user performance, it is still important to fully understand the reason behind the fundamental limitations of the ray casting interaction technique. This will also allow future interaction designers to come up with new ways to address the challenges.

2.2 3D Target Selection in VR

Stereo displays are beneficial for the selection of 3D targets in the near-field with ray casting [30, 44]. However, pointing throughput is typically below what users can achieve in 2D tasks [40, 43, 44].

Previous work has found that having targets at different depths affects performance negatively. For virtual ray pointing, Teather and Stuerzlinger [44] showed that varying target depth affects performance. Janzen et al. [26] found that pointing performance for targets at depths between 110 and 330 cm is affected. They also identified an effect of the user's distance to the screen, which would identify an issue related to the focal distance. For 3D virtual hand/wand pointing, Barrera and Stuerzlinger [3] found that lateral and depth movements were different when selecting targets displayed via a large stereo display. Batmaz et al. [5] verified that the same effect exists in current AR and VR headsets. All of these works did not conclusively identify the reason for the observed results.

One potential explanation for the lower performance with targets at varying depths is the presence of the VAC in current stereo display systems. As mentioned in the introduction, the VAC is caused by the way VR HDMs display 3D content, and it causes several problems in the human ocular system like 1) depth perception issues [13, 14], 2) visual fatigue [21, 22], due to the reduced stereo-acuity caused by the differences between focal and vergence distances [21], and 3) the eyes converge closer than required [23, 24]. All these issues affect the performance of the visual system [16, 47] and the cognitive load of the user [11]. Batmaz et al. [4] identified that the VAC affects the 3D selection of targets in peripersonal space with *virtual hand* interaction. However, no previous work has studied if the VAC affects target selection of distal targets using the ray casting selection technique.

2.3 Fitts' Law

Fitts' law [15] models human movement time (MT) for pointing, which is the time between initiation of the movement and the (successful) selection of the target. As ray casting involves (predominantly) rotational controller movements, we used Kopper et al.'s [30] formulation to model distal pointing task performance based on angular distances. See Equation 1. The constants a and b are empirical values, typically identified by linear regression. The logarithmic term represents the task difficulty and is called the index of difficulty (ID). In the calculation of angular ID, α represents the angular distance between targets and ω represents the angular target width. The constant k represents a relative weight between α and ω[30]:

$$MovementTime = MT = a + b \cdot \log_2(\frac{\alpha}{\omega^k} + 1) = a + b \cdot ID_{angular} \quad (1)$$

We also use throughput (THP) (based on effective measures) as defined in the ISO 9241-400:2015 document [25] (Equation 2):

$$THP = \frac{EffectiveIndexOfDifficulty}{MovementTime} = \frac{ID_e}{MT} \quad (2)$$

The effective index of difficulty (IDe) is defined by Equation 3. α_e represents the effective distance, e.g., the actual angular movement distance to the target position, and We is the effective target width, the distribution of selection coordinates, calculated as $\omega_e = 4.133 \times SD_x$. SD_x represents the distance between the selection points and the target center (projected onto the task axis), and this measure is useful to analyze the accuracy of the task performance [34, 35].

$$ID_e = \log_2(\frac{\alpha_e}{\omega_e^k} + 1) = \log_2(\frac{\alpha_e}{(4.133 \cdot SD_x)^k} + 1) \quad (3)$$

3 MOTIVATION & HYPOTHESES

Previous work examined the effects of VAC with virtual hand pointing, a commonly used interaction technique in VR [4]. Their results showed that the VAC has detrimental effects on the user performance in terms of time, error rate, and throughput. In this paper, we extend this work and examine the effect of the VAC on distal pointing with ray casting, which is the other most frequently used 3D interaction technique in VR. We expect to observe a decrease in user performance when the targets are away from the focal plane of the VR display, i.e., not in the focal plane.

We hypothesise (*H1*) that **user performance significantly decreases when the task is more complex**, i.e., when the user has to constantly adapt to combinations of accommodation and vergence cues that do not appear in the real world. In particular, we focus on situations where the user has to constantly changing their vergence onto targets at different visual depths, while keeping their accommodation constant. Previous work found that this behaviour causes visual fatigue [17]. Batmaz et al. [4] speculated that user performance degrades when the participant has to accomplish such a task, where the VAC varies. However, they did not study this effect.

We also hypothesize (*H2*) that **3D selection with ray casting will be negatively affected by the presence of VAC**. Previous work showed that ray casting methods are negatively affected by a change in target depth [18, 32]. Also, a stereo display with the VAC does exhibit a reduction in performance for 3D target selection with the virtual hand, as identified in previous work, e.g., by Barrera et al. [3] and Batmaz et al. [5]. Barrera et al. [3] identified that movements along the line of sight are ≈25% slower than movements in the lateral plane with current stereo displays. Similarly, in a comparison between both movement directions in AR and VR headsets, Batmaz et al.[5] also found a significant difference in throughput.

4 USER STUDY

To analyze the impact of the VAC on the user performance, we conducted an ISO 9241-411:2015 [25] multidirectional selection study with ray casting to targets at two different depth planes, one without a VAC, the other with a constant VAC. As the third condition, we forced participants to select targets at two different depth planes, i.e., induced a varying VAC.

4.1 Participants

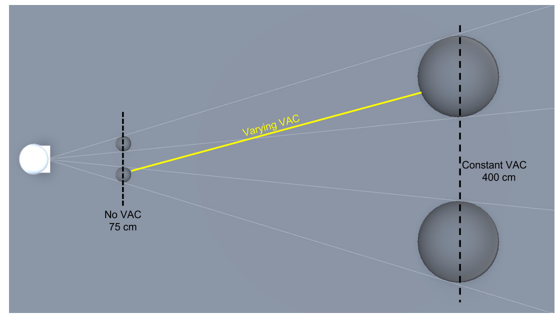
We recruited fifteen volunteer participants (7 female, 8 male) aged between 19 and 28 years (mean = 22.2, SD = 2.21). We recruited participants from different departments of the local university. Their participation was voluntary and no rewards were offered. None of the participants had prior experience with VR systems. Ten participants had normal and five corrected-to-normal vision. None of them reported color blindness or other visual impairments.

4.2 Apparatus

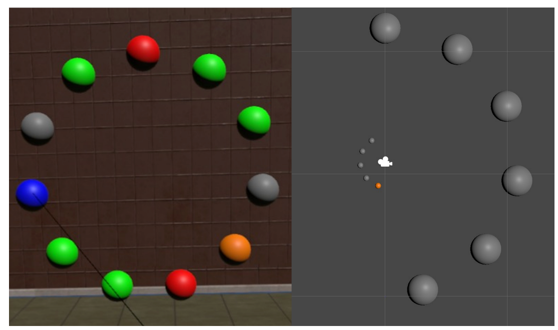
We conducted the experiment on an 11th Gen Intel(R) Core(TM) i7-11700F core 2.5 GHz, 32 GB RAM desktop PC with an NVIDIA GeForce RTX 3070 graphics card. We used an HTC VIVE Pro headset, one controller, and two 2.0 Lighthouse trackers. Furthermore, we designed and implemented the virtual environment in Unity3D version 2021.3.5f1.

4.3 Procedure

The participants started by filling out a consent form and a pre-experiment demographic survey. Then, we explained the experiment and conditions to them. The participants were instructed to use their dominant hand to hold the controller. For ray casting, we used a ray starting from the center of the controller. We showed a cursor sphere at the position where the ray collides with any of the objects in the scene. Participants selected targets with their non-dominant hand by pressing the space bar on the keyboard. We chose this selection method to eliminate the "Heisenberg Effect" [7].



(a)



(b)

Figure 1: Varying VAC condition. (a) The left image shows a top-view illustration with the viewer being at the left. Pairs of targets were either shown at 75 cm, without a VAC (left dashed line), at 400 cm, which induces a constant VAC (right dashed line), or with varying depth, i.e., with a varying VAC (yellow line). The white lines illustrate the projection of the targets. (b) On the left, we show the perspective view for varying VAC condition, where all targets appear to be the same size, even though they vary vastly in depth. On the right, we show an orthographic view of the same scene from the same camera position.

In the virtual environment, the participants were initially placed in an empty room. That virtual scene was similar to one used in previous work [5], but we used a different experimental task (and experimental design) in our current work. Then they were shown a circle of 11 equally distributed spherical targets in an arrangement that faces the participant centered in the middle of their view, i.e., a 3D configuration that (at first glance) looks like an ISO 9241-411 multidirectional selection task [25]. For the different conditions the depth of the targets varied, as described below. To investigate the VAC independent of perspective effects we varied the size of the target spheres so that further spheres appeared to subtend the same visual angle, i.e., all targets *appeared* to be the same size, but they were potentially at different distances, depending on the condition.

All spheres were grey at the beginning of each trial, except the target sphere, which was shown in orange. While the cursor collided with a sphere, we temporarily changed that sphere's color to blue for visual feedback, i.e., highlighted any intersected sphere [45]. The participant pointed with the VR controller to the target and pressed the space bar to select. If the cursor with collided (i.e., the ray intersected) the target sphere, we recorded the selection as a "hit" and changed the sphere color to green. Otherwise, if the cursor did not collide with the target sphere, then the selection was recorded as "miss", the sphere color was changed to red, and a short error sound plays as auditory error feedback [6].

The first target in each round was selected randomly from the 11 spheres. The direction of the next target was then selected to be either in a clockwise or anticlockwise manner directly across the circle of targets. Since they were eleven targets, the next target (orange) alternated across the circle till all targets were recorded as a "hit" or a "miss". We (re-)centered the circle before each round of trials to line up with the viewers eye height to ensure that the perspective cues lined up correctly every time.

The participants sat while they took part in the experiment, which limited their potential range of movement. We did not fix participants' head position, but we used the current participants' current head position and rotation after each circle of targets to align the targets, so that they appeared straight in front of the participant. So, even if participants moved, the system reset the position of the targets after (at most) 10 target selections.

In the experiment, we evaluated three different 3_{VAC} **VAC conditions**. The first condition was the **No VAC** condition, where the spheres where placed at the focal plane of the HTV Vive Pro, i.e., 75 cm from the viewer. The second was the **Constant VAC** condition, where the spheres were placed at 400 cm. The third condition was **Varying VAC** condition, where the spheres were placed at alternating depths (75 and 400 cm). In other words, if the first target appeared at 75 cm, then the next target appeared at 400 cm, and then the third at 75 cm, and so on. Due to the alternating depth of targets in the varying VAC condition, this means that half of the targets were close while the other half of the targets where far away, see the orthographic projection in Figure 1.

We deliberately chose the depths mentioned for the No Vac and Constant Vac conditions. In Batmaz et al. [4], the authors used three different depth distances, namely 40, 55, and 70 cm. To analyze the effect of the VAC on the multifocal and single focal displays, the authors positioned targets at 40 and 70 cm, which yields a difference of 1.08 diopters (i.e., 100/40 - 100/70 = 2.5 - 1.42 = 1.08). When we placed the targets at 75 cm and 400 cm for the study, this also corresponds to a difference of 1.08 diopters (i.e., 100/75 - 100/400 = 1.33 - 0.25). This enabled us to increase the comparability across the studies.

In Batmaz et al. [4] and Barrera et al. [3], the targets for the lateral movement conditions were placed 55 cm away from the participants. In the latter paper, the distance between targets was 30 cm and the authors used three different target sizes, 1.5, 2.5, and 3.5 cm [3]. To increase the comparability of our work, we used the same distances here. As both these works used only a single target distance we used *additional* target distances in this work, by varying the distances between targets ±5 cm, which widens the range of ID's our work explores relative to those previous works. We then converted the Euclidean distances and sizes to angular distances and sizes, based on the Kopper et al.'s previous work on angular measures [30] which yielded nine different IDs as shown in Table 1.

Table 1: Angular target size and target distance that are used in this work. The values marked with ∗ show values used in previous work [2].

Target Size (cm)	Target Distance (cm)	Angular Target Size (°)	Angular Target Distance(°)
1.5*	30*	1.45	30.51
2.5*	30*	2.42	30.51
3.5*	30*	3.39	30.51
1.5*	25	1.49	25.61
2.5*	25	2.48	25.61
3.5*	25	3.47	25.61
1.5*	35	1.42	35.30
2.5*	35	2.36	35.30
3.5*	35	3.31	35.30

Table 2: Data Analysis Results

	VAC	ID
Time	$F(2,28) = 92.5, p < 0.001, \eta^2 = 0.869$	$F(8,112) = 96.01, p<0.001, \eta^2 = 0.873$
Error rate	$F(2,28) = 5.73, p < 0.01, \eta^2 = 0.290$	$F(8, 112) = 3.18, p<0.001, \eta^2 = 0.189$
Throughput	$F(2,28) = 69.273, p < 0.001, \eta^2 = 0.832$	$F(8,112) = 20.307, p<0.001, \eta^2 = 0.592$
SD_x	$F(2,28) = 11.903, p<0.001, \eta^2 = 0.46$	$F(8,112) = 6.62, p<0.001, \eta^2 = 0.321$

The reason we used angular measures instead of Euclidean measures is to keep the (visually perceived) target sizes and distance the same regardless of which depth plane a target was positioned at. This normalizes the observed target size and distance by the participant regardless if the condition was **No VAC**, **Constant VAC**, or **Varying VAC** (or in other words, either 75 cm, 400 cm, or both), as shown in Figure 1b.

After the experiment, we asked the participants to fill a questionnaire regarding their insights and comments on the different VAC conditions.

4.4 Experimental Design

We used a within-subjects design with three **VAC conditions** (3_{VAC} = No VAC, Constant VAC, and Varying VAC). We measured task execution time (seconds), error rate (%), effective throughput (bits/s), and SD_x as recorded from the participant's trials. We counterbalanced the VAC conditions across participants with a Latin Square. We varied the angular index of difficulty $ID_{angular}$, by using three **angular target sizes** (3_{ATD}) and three **angular target distances** (3_{ATS}), which created 9 unique IDs. Angular target sizes and distances were randomly selected for each round of trials. Each subject performed $3_{VAC} \times 9_{ID} \times$ 11 repetitions = 297 trials.

5 RESULTS

We analyzed the data using Repeated Measures (RM) ANOVA in SPSS 24. We considered data as normally distributed when Skewness (S) and Kurtosis (K) were within ±1 [20, 36]. When the data was not normally distributed (even after a potential log transformation), we used ART [48]. We used the Bonferroni method for post-hoc analyses. Results are illustrated as means and standard error of means in figures.

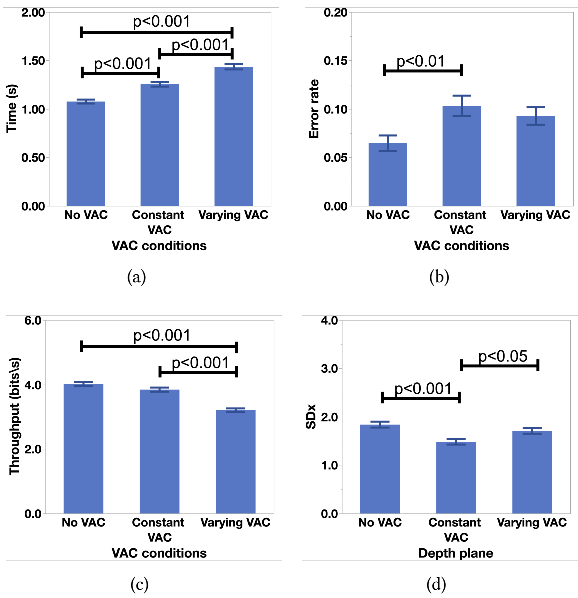


Figure 2: RM ANOVA results for (a) time, (b) error rate (c) throughput and (d) SD_x.

5.1 Time

Time was normally distributed after the log transformation (S = 0.89, K = 1.2). The results show that the VAC condition significantly affected time, Table 2. According to these results, it took longer to execute task with the varying VAC condition compared to the constant VAC condition, which in turn was also slower than the no VAC conditions (Figure 2(a)). ID results were also significant, Table 2.

5.2 Error rate

The error rate was not normally distributed, even after a log transformation (S = 1.55, K = 2.87). The ART results identify that there is a significant difference for VAC conditions, Table 2. As shown in Figure 2(b), participants' error rate increased with the constant VAC condition compared to the no VAC condition. ID results were also significant, Table 2.

5.3 Throughput

Throughput data was normally distributed (S = 0.46 and K = 0.66). The results in Figure 2 (c) illustrate that the varying VAC condition exhibited significantly reduced throughput performance compared to the other two conditions, Table 2. ID results were also significant, Table 2.

5.4 *SDx*

SD_x data was normally distributed after log-transformation (S= 0.18, K = -0.19). We found a significant difference for VAC conditions,Table 2. The results in Figure 2 (d) highlight that accuracy of the participants significantly increased for the constant VAC condition. ID results were also significant, Table 2.

5.5 Fitts' Law Results

Similar to previous work [30], we found the highest R^2 value when k is 1 (R^2=0.96), with MT = -0.11 + 0.36 * ID. In separate analysis for each condition, we found that the no VAC condition can be modeled as MT = -0.30 + 0.36 * ID, R^2 =0.98, the constant VAC condition as MT = -0.01 + 0.33 * ID, R^2 =0.9, and varying VAC as MT = -0.032 + 0.38 * ID, R^2 =0.92. These results are shown in Figure 3.

5.6 Questionnaire Results

After the experiment, we used a short questionnaire to collect feedback from the participants. We asked participants which experimental condition they preferred and about their reasoning behind their preference. 8 participants preferred No VAC, and 7 preferred the Varying VAC condition, while none liked the Constant VAC condition. Participants commented on No VAC: *"it requires me to move my hand less," "I could immediately reach the target. It was very fast," and "it was easier for me to choose the targets when they where close."* They also commented on the Varying VAC: *"It was more fun", "changes your attention" and "it was interesting to do. Like a game. It was balanced between close and far".*

We also asked the participants if it was easy to select targets with the different VAC conditions. According to our 7-point Likert scale results (1-I totally disagree, 7-I totally agree), participants agreed that No VAC was the easier task (mean: 5.86, median: 6, standard deviation: 1.41). They agreed less to the statement that it was easy to select targets the in Constant VAC condition (mean: 4.4, median: 5, standard deviation: 1.45). As for the Varying VAC condition, there was a slightly higher agreement compared to Constant VAC (mean: 4.73, median: 5, standard deviation: 1.53).

Finally, we asked participants if they felt fatigue after the experiment (1- I totally feel normal 7- I feel completely fatigue). Participants neither felt strong physical fatigue (mean: 2.13, median: 2, standard deviation: 1.35) nor mental fatigue (mean: 1.26, median: 1, standard deviation: 0.59).

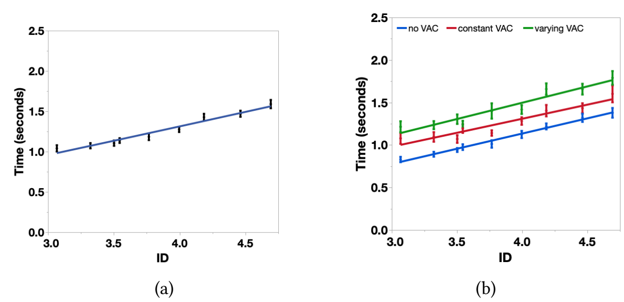


Figure 3: ID analysis for (a) whole user study, and for (b) each VAC conditions.

6 DISCUSSION

In this work, we analyzed the effect of the VAC on distal pointing, by positioning targets either in configurations without a VAC, a constant VAC, or a varying VAC. While we used a visually somewhat similar virtual scene, the experimental design and the experimental task were both substantially different from previous work [5].

The results showed that participants were faster, made fewer errors, and their throughput increased when there was no VAC, i.e., at the focal plane of the used VR HMD. On the other hand, when we placed the targets at a depth plane with a constant VAC (at 400 cm), we observed a significant user performance decrease for time, error rate, and accuracy. Further, when we varied the VAC by placing target at two different target planes (one with and one without a VAC), the user performance decreased even more for time, error rate, and throughput, and exhibited a significant difference even to the constant VAC condition. These results show that for ray casting user performance decreases significantly in conditions when there is VAC present, which supports our *H1* that **user performance significantly decreases when the task gets more complicated** such as where the user has to focus on targets at different visual depths. Moreover, the study presented here support Batmaz et al.'s [4] speculation that user performance further decreases when the VAC varies during a task.

Our results here also extend previous findings on the effect of the VAC on virtual hand pointing [4], which indicate that the VAC likely affects most (if not all) 3D selection techniques. Our results support *H2* that **3D selection with ray casting will be negatively affected by the presence of VAC.** It is interesting to observe that the condition without the VAC exhibits a throughput of ≈4 bps, which is in the performance range frequently observed for mouse pointing [40]. In other words, the pointing performance for targets with no VAC approaches (or matches) that of a mouse. The introduction of a VAC – either a constant VAC or, worse yet, a varying VAC – reduces the performance up to ≈ 25% in terms of time and throughput, which coincides with the differences observed in previous work that used a different methodology [3].

We chose 75 cm and 400 cm depth distances to match the change of 1.08 diopters used in Batmaz et al. [4]. This decision allows us to extend their findings for distances beyond arm's reach with a different interaction modality, raycasting. After all, we believe that visually guided movements, such as the pointing movements we are investigating here, are influenced by issues in depth perception. We used a depth distance of 75 cm, as this is the focal plane in at least some current HMD's, i.e., the distance where no VAC occurs. Batmaz et al. [4] had already investigated a virtual hand condition, but with much closer targets (less than 75 cm). To simplify the experimental design while still enabling comparability, we focused only on a single depth change with the same change in diopters as Batmaz et al. [4].

We did not change the stimuli, beyond the change necessary to induce the VAC in the experiment. The color of the target sphere, selection color, error sounds, and all other aspects were the same. We only changed the distance of the spheres and proportionally matched their size so that participants perceived them to have the same (angular) target size in all conditions, making them equally easy to select with raycasting. Participants were not informed that we manipulated the targets so that they appeared to have the same size at different distances. Further, we did not specify which experimental condition a user would experience at the beginning of each sequence of trials. Thus, we believe that participants were not aware of the precise nature of our experimental task design.

Previous work on stereo displays [44] had already shown that for raycasting the execution time and throughput of participants who perceive the same object size do not significantly change, regardless of target distance. After all, raycasting relies (predominantly) only on 2DOF rotation to move the ray/cursor to select objects. Thus, even when the targets were further away, participants only needed to rotate the VR controller by the same amount for all conditions, and the observed differences are thus directly attributable to the VAC. Another example of work that might have been affected by the presence of the VAC is Janzen et al. [26]. In their work, selection performance changed across screen and target depths between 110 and 330 cm. Yet, their conditions that match the target distance to the screen distance behaved differently than other conditions. Our results show that the absence of a VAC helps user performance, which might explain the difference they observed. However, this hypothesis needs to be verified in future work.

The questionnaire results also reveal that participants seem to be easily able to perceive the difference between the condition without a VAC and conditions that had a VAC and comments and ratings speak to the no VAC condition being easiest to use. Interestingly, some participants found the Varying VAC condition to be like a game, i.e., they recognized that successful and fast pointing to the targets involved a noticeable challenge. This may be a side-effect of our participant pool with mostly young adults. Thus our study should be repeated with a wider population. In particular, it would be also appropriate to include professionals that are using or plan to use VR HMD's in their work.

Our results exhibit high effect sizes (η^2>.14), which means the statistical results are very likely to be robust and replicable. Following advice on experimental design [33], we also aimed to keep the number of participants comparable with related work. Previous work on the VAC [4] conducted their user study with 18 participants. Considering all constraints and to keep everything counter-balanced, we chose to use 15 participants.

Overall, given our results, we suggest that practitioners, engineers, and 3D user interface designers should not vary the target depth for pointing task to optimize the user's pointing performance. This means that environments should be designed so that targets are placed at (roughly) the same depth distance from the user for ray casting.

6.1 Limitations

Even though we conducted our experiment using a HTC Vive Pro, the results of our experiment must be replicated in other headsets to see if the generalize to all VR HMDs. After all, and technically speaking, the empirical results in this paper only demonstrate a performance decrease in an HTC Vive Pro HMD. Having said this, and given the results of previous work [3] also demonstrated an effect of the VAC in a stereo display system without lenses, it seems likely that our results will hold in other headsets and other stereo display systems, too.

Given that our results identify that the VAC is the main cause for the observed performance differences, we believe that the results can also be explained in terms of the change in the focal plane, i.e., in terms of diopters. In this study we only analyzed the VAC for targets with a difference of 1.08 diopters (i.e., 100/75 - 100/400 = 1.33 - 0.25). We used depth distances of 75 and 400 cm to elicit the VAC with the same change in diopters as in [4], We also used HTC Vive

Pro since (as far as we can tell) the focal point of this device is at 75 cm. The Oculus Quest 2's [37] focal point is at 1.3 m, and using it for this first study would have made it more challenging to directly extend the results of previous work [4]. We will still acknowledge that we used a single HMD and a single change in depth distance and that future work should investigate if our results hold for other HMDs and other changes in depth distance.

Another limitation of our study concerns the Fresnel lenses that are used by HTC Vive Pro. These lenses distort the image coming to the user which means that the focal plane may not be perfectly flat. Thus, we can only claim that the targets we used in this study were placed "roughly" at the focal plane of the VR headset.

Batmaz et al.'s [5] previous work on the effect of stereo display cue conflicts identified that there is no user performance difference in terms of time, error rate, and throughput between VR and AR headsets. Both systems exhibit the same differences due to the VAC. Since many commercial AR-headsets use lenses, we speculate also that our results will extend to AR headsets. Yet, these claims need to be further investigated.

7 CONCLUSION

In this work, we investigated the effect of the VAC on distal pointing with ray casting. The results show that user performance in terms of time and throughput significantly decreases when the targets are not positioned at the focal plane of the VR HMDs. Furthermore, when users have to select consecutive targets at different depth planes, their throughput performance decreases even more. Thus, we suggest that practitioners, developers, and designers place target at the focal point of (single-focal) VR HMDs for the best pointing performance. Alternatively, multi-focal VR HMD's should be considered, although we point out that we still need to verify if the absence of the VAC in multi-focal HMD's removes the observed drop in pointing performance.

In the future, we plan to extend our work to other VR and AR HMDs to quantify the effect of VAC for distal pointing. We also want to investigate VAC in other distal pointing methods, such as with the HOMER technique. Similarly, we plan to evaluate the effects of the VAC for movements with different changes of distance and examine the impact of the outcomes on 3D UI design.

REFERENCES

[1] Marc Baloup, Thomas Pietrzak, and Géry Casiez. 2019. RayCursor: A 3D Pointing Facilitation Technique Based on Raycasting. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3290605.3300331

[2] Mayra Donaji Barrera Machuca and Wolfgang Stuerzlinger. 2018. Do Stereo Display Deficiencies Affect 3D Pointing?. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI EA '18)*. Association for Computing Machinery, New York, NY, USA, 1–6. https://doi.org/10.1145/3170427.3188540

[3] Mayra Donaji Barrera Machuca and Wolfgang Stuerzlinger. 2019. The Effect of Stereo Display Deficiencies on Virtual Hand Pointing. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3290605.3300437

[4] Anil Ufuk Batmaz, Mayra Donaji Barrera Machuca, Junwei Sun, and Wolfgang Stuerzlinger. 2022. The Effect of the Vergence-Accommodation Conflict on Virtual Hand Pointing in Immersive Displays. In *CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 633, 15 pages.

[5] Anil Ufuk Batmaz, Mayra Donaji Barrera Machuca, Duc Minh Pham, and Wolfgang Stuerzlinger. 2019. Do Head-Mounted Display Stereo Deficiencies Affect 3D Pointing Tasks in AR and VR?. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 585–592. https://doi.org/10.1109/VR.2019.8797975

[6] Anil Ufuk Batmaz and Wolfgang Stuerzlinger. 2021. The Effect of Pitch in Auditory Error Feedback for Fitts' Tasks in Virtual Reality Training Systems. In *Conference on Virtual Reality and 3D User Interfaces (VR '21)*. 85–94. https://doi.org/10.1109/VR50410.2021.00029

[7] Doug Bowman, Chadwick Wingrave, Joshua Campbell, and Vinh Q Ly. 2001. Using pinch gloves (tm) for both natural and abstract interaction techniques in virtual environments.

[8] Doug A. Bowman and Larry F. Hodges. 1997. An Evaluation of Techniques for Grabbing and Manipulating Remote Objects in Immersive Virtual Environments. In *Proceedings of the 1997 Symposium on Interactive 3D Graphics* (Providence, Rhode Island, USA) *(I3D '97)*. Association for Computing Machinery, New York, NY, USA, 35–ff. https://doi.org/10.1145/253284.253301

[9] Yuan Chen, Junwei Sun, Qiang Xu, Edward Lank, Pourang Irani, and Wei Li. 2021. Global Scene Filtering, Exploration, and Pointing in Occluded Virtual Space. In *Human-Computer Interaction – INTERACT 2021 - 18th IFIP TC 13 International Conference, Proceedings (Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics))*, Carmelo Ardito, Rosa Lanzilotti, Alessio Malizia, Alessio Malizia, Helen Petrie, Antonio Piccinno, Giuseppe Desolda, and Kori Inkpen (Eds.). Springer Science and Business Media Deutschland GmbH, Germany, 156–176. https://doi.org/10.1007/978-3-030-85607-6_11

[10] J. Clement. 2022. Virtual reality (VR) gaming revenue worldwide from 2017 to 2024.

[11] François Daniel and Zoï Kapoula. 2019. Induced vergence-accommodation conflict reduces cognitive performance in the Stroop test. *Scientific Reports* 9, 1 (2019), 1–13. https://doi.org/10.1038/s41598-018-37778-y

[12] Pham Duc-Minh and Wolfgang Stuerzlinger. 2019. Is the Pen Mightier than the Controller? A Comparison of Input Devices for Selection in Virtual and Augmented Reality. In *25th Symposium on Virtual Reality Software and Technology (VRST '19)*. ACM, Article 35, 11 pages. https://doi.org/10.1145/3359996.3364264

[13] Frank H. Durgin, Dennis R. Proffitt, Thomas J. Olson, and Karen S. Reinke. 1995. Comparing depth from motion with depth from binocular disparity. *Journal of Experimental Psychology: Human Perception and Performance* 21, 3 (1995), 679–699. https://doi.org/10.1037/0096-1523.21.3.679

[14] G. N. Dutton, A. Saaed, B. Fahad, R. Fraser, G. McDaid, J. McDade, A. Mackintosh, T. Rane, and K. Spowart. 2004. Association of binocular lower visual field impairment, impaired simultaneous perception, disordered visually guided motion and inaccurate saccades in children with cerebral visual dysfunction a retrospective observational study. *Eye* 18, 1 (jan 2004), 27–34. https://doi.org/10.1038/sj.eye.6700541

[15] 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.

[16] Tetsuya Fukushima, Masahito Torii, Kazuhiko Ukai, James S. Wolffsohn, and Bernard Gilmartin. 2009. The relationship between CA/C ratio and individual differences in dynamic accommodative responses while viewing stereoscopic images. *Journal of Vision* 9, 13 (12 2009), 21–21. https://doi.org/10.1167/9.13.21 arXiv:https://arvojournals.org/arvo/content_public/journal/jov/933531/jov-9-13-21.pdf

[17] Joseph L. Gabbard, Divya Gupta Mehra, and J. Edward Swan. 2019. Effects of AR Display Context Switching and Focal Distance Switching on Human Performance. *IEEE Transactions on Visualization and Computer Graphics* 25, 6 (2019), 2228–2241. https://doi.org/10.1109/TVCG.2018.2832633

[18] Evan D. Graham and Christine L. MacKenzie. 1996. Physical versus virtual pointing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '96)*. ACM Press, New York, New York, USA, 292–299. https://doi.org/10.1145/238386.238532

[19] Tovi Grossman and Ravin Balakrishnan. 2005. The bubble cursor: enhancing target acquisition by dynamic resizing of the cursor's activation area. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 281–290.

[20] Joseph F Hair Jr, William C Black, Barry J Babin, and Rolph E. Anderson. 2014. *Multivariate data analysis.* Pearson Education Limited.

[21] 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 (mar 2008), 33.1–30. https://doi.org/10.1167/8.3.33

[22] Hyungki Hong and Seok Hyon Kang. 2015. Measurement of the lens accommodation in viewing stereoscopic displays. *Journal of the Society for Information Display* 23, 1 (jan 2015), 19–26. https://doi.org/10.1002/jsid.303

[23] Anke Huckauf, Mario H. Urbina, Irina Böckelmann, Lutz Schega, Rüdiger Mecke, Jens Grubert, Fabian Doil, and Johannes Tümler. 2010. Perceptual issues in optical-see-through displays. *Proceedings - APGV 2010: Symposium on Applied Perception in Graphics and Visualization* 1, 212 (2010), 41–48. https://doi.org/10.1145/1836248.1836255

[24] J. Iskander, Mohammed Hossny, and S. Nahavandi. 2019. Using biomechanics to investigate the effect of VR on eye vergence system. *Applied Ergonomics* 81, August 2018 (2019), 102883. https://doi.org/10.1016/j.apergo.2019.102883

[25] ISO. 2015. ISO 9241-400:2015 Ergonomics of human-system interaction - Part 411: Evaluation methods for the design of physical input devices.

[26] Izabelle Janzen, Vasanth K. 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, California, USA) *(CHI '16)*. ACM, New York, NY, USA, 188–199. https://doi.org/10.1145/2858036.2858244

[27] Nicholas Katzakis, Kiyoshi Kiyokawa, and Haruo Takemura. 2013. Plane-Casting: 3D Cursor Control With A SmartPhone. In *Proceedings of the 11th Asia Pacific Conference on Computer Human Interaction* (Bangalore, India) *(APCHI '13)*. Association for Computing Machinery, New York, NY, USA, 199–200. https://doi.org/10.1145/2525194.2525275

[28] N. Katzakis, R. J. Teather, K. Kiyokawa, and H. Takemura. 2015. INSPECT: Extending plane-casting for 6-DOF control. In *2015 IEEE Symposium on 3D User Interfaces (3DUI)*. IEEE Computer Society, Los Alamitos, CA, USA, 165–166. https://doi.org/10.1109/3DUI.2015.7131752

[29] Joohwan Kim, David Kane, and Martin S. Banks. 2014. The rate of change of vergence-accommodation conflict affects visual discomfort. *Vision Research* 105 (2014), 159–165. https://doi.org/10.1016/j.visres.2014.10.021 arXiv:NIHMS150003

[30] 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.

[31] Gregory Kramida. 2016. Resolving the vergence-accommodation conflict in head-mounted displays. *IEEE Transactions on Visualization and Computer Graphics* 22, 7 (2016), 1912–1931. https://doi.org/10.1109/TVCG.2015.2473855 arXiv:arXiv:1011.1669v3

[32] Chiuhsiang J. Lin and Bereket H. Woldegiorgis. 2017. Egocentric distance perception and performance of direct pointing in stereoscopic displays. *Applied Ergonomics* 64 (2017), 66 – 74. https://doi.org/10.1016/j.apergo.2017.05.007

[33] I Scott MacKenzie. 2013. *Human-computer interaction: An empirical research perspective*. Morgan Kaufmann.

[34] I. Scott MacKenzie and Poika Isokoski. 2008. Fitts' Throughput and the Speed-Accuracy Tradeoff. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Florence, Italy) *(CHI '08)*. Association for Computing Machinery, New York, NY, USA, 1633–1636. https://doi.org/10.1145/1357054.1357308

[35] I Scott MacKenzie and Aleks Oniszczak. 1998. A comparison of three selection techniques for touchpads. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM Press/Addison-Wesley Publishing Co., 336–343.

[36] Paul Mallery and Darren George. 2003. *SPSS for Windows step by step: a simple guide and reference*. Pearson.

[37] Meta. 2022. Oculus Quest 2.

[38] Moaaz Hudhud Mughrabi, Aunnoy K Mutasim, Wolfgang Stuerzlinger, and Anil Ufuk Batmaz. 2022. My Eyes Hurt: Effects of Jitter in 3D Gaze Tracking. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. 310–315. https://doi.org/10.1109/VRW55335.2022.00070

[39] Pimax. 2019. Pimax.

[40] R. William Soukoreff and I. Scott MacKenzie. 2004. Towards a Standard for Pointing Device Evaluation, Perspectives on 27 Years of Fitts' Law Research in HCI. *Int. J. Hum.-Comput. Stud.* 61, 6 (Dec. 2004), 751–789.

[41] Helmut Strasser and Karl-Werner Müller. 1999. Favorable movements of the hand-arm system in the horizontal plane assessed by electromyographic investigations and subjective rating. *International Journal of Industrial Ergonomics* 23, 4 (1999), 339–347. https://doi.org/10.1016/S0169-8141(98)00050-X

[42] Junwei Sun, Wolfgang Stuerzlinger, and Dmitri Shuralyov. 2016. SHIFT-Sliding and DEPTH-POP for 3D Positioning. In *Proceedings of the 2016 Symposium on Spatial User Interaction* (Tokyo, Japan) *(SUI '16)*. Association for Computing Machinery, New York, NY, USA, 69–78. https://doi.org/10.1145/2983310.2985748

[43] Robert J Teather, Andriy Pavlovych, Wolfgang Stuerzlinger, and I Scott MacKenzie. 2009. Effects of tracking technology, latency, and spatial jitter on object movement. In *3D User Interfaces, 2009. 3DUI 2009. IEEE Symposium on*. IEEE, 43–50.

[44] R. J. Teather and W. Stuerzlinger. 2011. Pointing at 3D targets in a stereo head-tracked virtual environment. In *2011 IEEE Symposium on 3D User Interfaces (3DUI)*. 87–94.

[45] Robert J Teather and Wolfgang Stuerzlinger. 2014. Visual aids in 3D point selection experiments. In *Proceedings of the 2nd ACM symposium on Spatial user interaction*. ACM, 127–136.

[46] Varjo. 2021. VR-3.

[47] Cyril Vienne, Laurent Sorin, Laurent Blondé, Quan Huynh-Thu, and Pascal Mamassian. 2014. Effect of the accommodation-vergence conflict on vergence eye movements. *Vision Research* 100 (2014), 124–133. https://doi.org/10.1016/j.visres.2014.04.017

[48] Jacob O. Wobbrock, Leah Findlater, Darren Gergle, and James J. Higgins. 2011. The Aligned Rank Transform for Nonparametric Factorial Analyses Using Only ANOVA Procedures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada) *(CHI '11)*. ACM, New York, NY, USA, 143–146. https://doi.org/10.1145/1978942.1978963