Measuring the Effect of Stereo Deficiencies on Peripersonal Space Pointing

Anil Ufuk Batmaz * Concordia University
Moaaz Hudhud Mughribi † Kadir Has University
Mine Sarac ‡ Kadir Has University
Mayra Barrera Machuca § Dalhousie University
Wolfgang Stuerzlinger ¶ Simon Fraser University

ABSTRACT

State-of-the-art Virtual Reality (VR) and Augmented Reality (AR) headsets rely on singlefocal stereo displays. For objects away from the focal plane, such displays create a vergence-accommodation conflict (VAC), potentially degrading user interaction performance. In this paper, we study how the VAC affects pointing at targets within arm's reach with virtual hand and raycasting interaction in current stereo display systems. We use a previously proposed experimental methodology that extends the ISO 9241-411:2015 multi-directional selection task to enable fair comparisons between selecting targets in different display conditions. We conducted a user study with eighteen participants and the results indicate that participants were faster and had higher throughput in the constant VAC condition with the virtual hand. We hope that our results enable designers to choose more efficient interaction methods in virtual environments.

Index Terms: Human-centered computing—Human Computer Interaction (HCI); Human-centered computing—Virtual Reality; Human-centered computing—Pointing

1 INTRODUCTION

Modern head-mounted displays (HMDs) are self-contained, wireless, lightweight, and capable of creating high-definition graphics for virtual reality (VR), e.g., Quest 2 or Pico 4, or augmented reality (AR), e.g., HoloLens 2 or Magic Leap 2. These systems are also affordable and accessible to a wide audience. Other commercial HMDs offer 4K resolution with low latency 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 and the Pimax 4K. Thanks to these technological advances, it is expected that between 2019 and 2024, there will be 34 million VR HMDs sold worldwide [49].

Despite the popularity and advantages of modern HMDs, users still experience challenges while interacting with targets positioned at varying distances [3, 6]. Barrera and Stuerzlinger [4] and Batmaz et al. [8] hypothesized that these challenges are likely to be caused by the way current stereo displays render content, which creates a mismatch between focusing the eyes on the display plane (accommodation) and rotating the eyes to see the content at its correct visual depth (vergence). This effect is called the vergence-accommodation conflict (VAC), and it does not occur for targets in the real world or virtual targets on the focal plane of display systems.

To verify the hypothesis that the VAC affects 3D pointing performance, Batmaz et al. [6] built a custom-made multifocal display to study the effect of VAC on virtual hand 3D selection through a comparison of virtual hand pointing in a multifocal and singlefocal stereo display. The results confirmed that their performance was worse in the singlefocal condition, which suffered from the VAC. However, their study used a very simple pointing task that involved only motions with and without a change in visual depth. To analyze this further, Batmaz et al. [7] proposed a new experimental methodology to demonstrate the (isolated) effect of the VAC on 3D selection within a singlefocal display system, with raycasting for targets beyond arm's reach. The results showed that a varying VAC condition significantly increases the required time and decreases the throughput performance.

In this work, we aim to understand the impact of the VAC within peri-personal space, i.e., within arm's reach, for general pointing tasks in a VR HMD. We hypothesize (*H*) that **3D selection with raycasting and virtual hand is negatively affected by the presence of the VAC**. Our work extends previous work on the impact of the VAC [4, 6] by including movements that require both lateral and visual depth changes, which better encompasses general 3D user interface interaction tasks. In such scenarios, users regularly use "diagonal" movements and can also rely on other depth cues besides vergence and accommodation to identify the target position.

To achieve our goal, we conducted an experiment with 18 participants based on Fitts' law. Building on Batmaz et al.'s innovative experimental methodology for distal pointing [7], we study the selection of targets within arm's reach, i.e., targets in front of the focal plane. The results of the study presented here contradict the results of previous work on the effect of the VAC. For the selection of targets at different depths in peri-personal space, our results do not reveal a detrimental effect of the VAC. Similarly, having to regularly compensate for the VAC did not decrease the user performance in terms of time, error rate, nor throughput. We speculate that our results are an outcome of bio-mechanical limitations, the distortions in the Fresnel lenses used in commercial HMDs, and/or the impact of other depth cues in the environment.

In summary, our contributions are the following:

- We extend the novel 3D pointing evaluation methodology proposed by Batmaz et al. [7] to enable its use for targets within arm's reach.
- We demonstrate the effect of the VAC in a singlefocal display system for targets within arm's reach using virtual hand and raycasting interactions.
- We quantify the effect of the VAC on the selection of targets within arm's reach in a commercial VR HMD.

2 RELATED WORK

In this section, we discuss the interaction techniques available for peri-personal space 3D pointing. We also discuss previous work on the effect of VAC on 3D pointing.

2.1 Fitts' Law

Fitts' law [22] models human movement time (MT) for pointing, which is the time between the initiation of the movement and the

*e-mail: ufuk.batmaz@concordia.ca
†e-mail: moaaz.mazen@stu.khas.edu.tr
‡e-mail: mine.sarac@khas.edu.tr
§e-mail: mbarrera@dal.ca
¶e-mail: w.s@sfu.ca

(successful) selection of the target. Although the original formulation did not differentiate between interaction techniques, previous work found that the difference in hand movements between raycasting and virtual hand requires the use of different formulations [31]. Next, we present each of these formulations:

2.1.1 Shannon Formulation

Previous work has proposed different Fitts' law formulations for 3D pointing movements using the virtual hand [3, 15, 16, 40], but no standardized formulation exists. Thus, we used the Shannon-Formulation (Equation 1) proposed by Mackenzie et al. [33] to calculate MT for 3D [47,48]:

$$MT = a + b \cdot \log_2\left(\frac{D}{W} + 1\right) = a + b \cdot ID \tag{1}$$

In the above equation D and W are the target distance and size, respectively, while a and b are empirically derived via linear regression. The logarithmic term in Fitts' law is known as the index of difficulty (ID) and indicates the overall pointing task difficulty. We also use throughput (THP) based on effective measures as defined in the ISO 9241-400:2015 document [29] (Equation 2):

$$THP = \frac{EffectiveIndexOfDifficulty}{MovementTime} = \frac{ID_e}{MT} \tag{2}$$

The effective index of difficulty (IDe) is defined by Equation 3, where A_e is the amplitude of the movement and W_e is the effective target width. W_e is related to the standard deviation between the selection position and the target center (SD_x), and this measure is useful to analyze the accuracy of the task performance [34, 35]:

$$ID_e = \log_2\left(\frac{A_e}{W_e} + 1\right) = \log_2\left(\frac{A_e}{(4.133 \cdot SD_x)} + 1\right) \tag{3}$$

2.1.2 Angular Fitts Law

Kopper et al.'s [31] formulation applies Fitts' law to rotational control movements. This formulation is used for raycasting, as its performance is determined by angular distances (See Equation 4). The difference relative to Equation 1 is the calculation of angular ID (ID_A), where α represents the angular distance between targets and ω the angular target width. The constant k represents a relative weight between α and ω [31], which is typically set to 1:

$$MT = a + b \cdot \log_2\left(\frac{\alpha}{\omega^k} + 1\right) = a + b \cdot ID_A \tag{4}$$

When calculating THP for angular movements, the angular distances simply replace the Euclidean ones in the IDe equation. In Equation 5, α_e represents the effective angular distance, e.g., the actual angular movement distance to the target position, and ω_e is the effective angular target width, the distribution of the angular selection coordinates, calculated as ω_e = 4.133×SD_x. SD_x represents the angular distance between the selection points and the target center (projected onto the angular task axis).

$$ID_e = \log_2\left(\frac{\alpha_e}{\omega_e^k} + 1\right) = \log_2\left(\frac{\alpha_e}{(4.133 \cdot SD_x)^k} + 1\right) \tag{5}$$

2.2 Peri-personal Space 3D Pointing

3D pointing in peri-personal space is the selection of targets close to the user, within arm's length, up to about 75 cm away. Two of the most popular selection methods for nearby targets are raycasting and the virtual hand technique.

The *virtual hand technique* allows users to select a target by intersecting it with their hand or an input device such as a controller, and then pressing a button to select it. The main limitation of the virtual hand technique is that targets need to be within arms' reach. Users also need to accurately perceive the position of the target and then move their hand there, which might be affected by depth perception issues [3, 8]. Finally, the absence of haptic feedback when selecting an object might also affect user performance [48].

The *raycasting technique* is used to select distal targets, but new systems also rely on it to select nearby objects like menus on the opposite hand as is an easy-to-understand technique that permits accurate selection at shorter distances. One of the main limitations of raycasting is that it is subject to the detrimental effects of unintentional hand tremor and/or tracker orientation variations, i.e., jitter, which affects the precision of the selection [10, 27, 47]. Another limitation is that targets that are very close to the user can require large angular movements [31].

When comparing user performance between virtual hand and raycasting, Teather and Stuerzlinger [48] found that raycasting has worse performance than the virtual hand, as shown by a slower selection time, higher error rate, and smaller throughput. They hypothesized that raycasting is more susceptible to tracker jitter amplification, but did not focus on other causes like the presence of the VAC. In summary, to better understand the reasons behind the limitations of both interaction techniques it is important to identify the effect of the VAC for 3D pointing in peri-personal space. This will also allow future interaction designers to come up with new ways to address these challenges, especially since techniques that mix virtual hand and raycasting are becoming more popular [46].

2.3 Effect of VAC on 3D pointing

Previous work found that 3D selection of targets benefits from stereo displays [31, 48], but that pointing throughput performance is lower than what users can achieve in 2D tasks [45, 47, 48].

For selection of 3D targets in peri-personal space with stereo displays, a change in depth between targets is a likely reason for this lower performance. When using a virtual hand, execution time and throughput are lower for 3D movements in visual depth compared to lateral movements. For example, and in a comparison with a real-world setup, Barrera and Stuerzlinger [3] found that lateral and depth movements were different when selecting targets shown on a large stereo display. Batmaz et al. [8] verified that the same effect exists in current AR and VR HMDs. We could not find previous work that evaluated the effect of depth on raycasting interaction for nearby targets. For distal pointing, Teather and Stuerzlinger [48] showed that varying target depth affects performance. Janzen et al. [30] also found that pointing performance is affected for targets at depths between 110 and 330 cm. They also identified an effect of the user's distance to the screen, which could indicate an issue related to the focal distance.

When selecting nearby targets, users use nonpictorial depth cues like stereopsis, motion parallax, and convergence and accommodation [14, 19, 42, 43]. Here we focus on the VAC, which is caused by the way VR HMDs display 3D content. Problems caused by the VAC in the human ocular system include: 1) depth perception issues [20, 21], 2) visual fatigue due to the reduced stereo-acuity caused by the differences between focal and vergence distances [25], and 3) that the eyes converge closer than required [26, 28]. All these issues affect the performance of the visual system [23, 50] and increase the cognitive load of the user [17]. Batmaz et al. [7] identified an effect of the VAC on the selection of distal 3D targets using *raycasting*. Their results show that when selecting targets at the HMDs' focal plane, e.g., in a condition without the VAC, participants exhibited better performance than when there was a constant or varying VAC. Participants were faster, made fewer errors, and showed higher throughput. However, this work did not study nearby targets within arms' reach. Previously, Batmaz et al. [6] identified that the VAC affects the 3D selection of targets in peri-personal space with *virtual hand* interaction, albeit only for ideal motions with only lateral or only depth components. They used a custom-made stereo display but did not study if the VAC affects target selection of nearby targets with the raycasting selection technique.

3 User Study

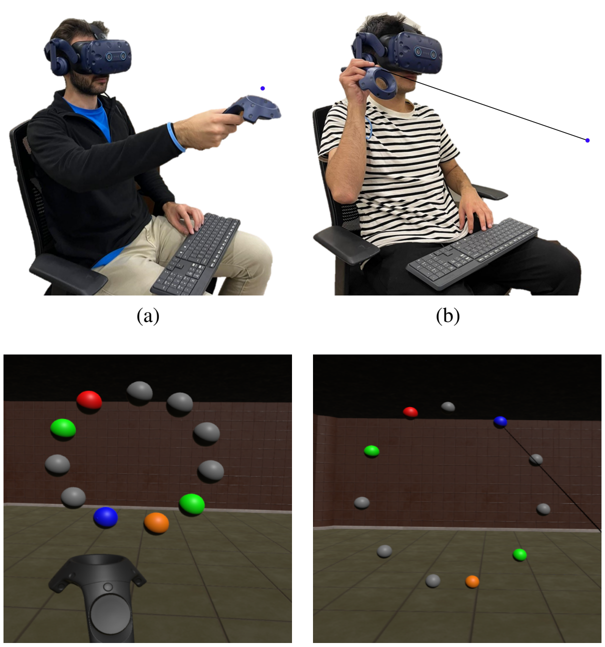
3.1 Participants

We recruited eighteen participants from the local university, five male and thirteen female with ages ranging between 18 and 33 (mean 21.61, stddev 3.18). The participants received no compensation for the experiments. Eight participants had normal vision, the other ten had corrected-to-normal vision. No-one reported color vision deficiencies. Regarding prior experience with VR, four participants had none, five had experienced it 1-3 times, three 3-5 times, and the remaining six had experienced it 5 or more times.

3.2 Apparatus

We conducted the experiment on an 11$^{\text{th}}$ Gen Intel(R) Core(TM) i7-11700F at 2.5 GHz, 32 GB RAM desktop PC with an NVIDIA GeForce RTX 3070 graphics card. We used an HTC VIVE Pro HMDs, with one controller and two 2.0 Lighthouse base stations. Further, we used Unity3D version 2021.3.5f1 to design and implement the virtual environment.

3.3 Procedure



(a) (b) (c) (d)

Figure 1: Experiment setup: Participants hold the controller differently in (a) the virtual hand condition and (b) the raycasting condition. Both images also show how the cursor/ray is represented in the virtual environment (VE), either as a cursor above the controller or as a ray. (c) Virtual hand condition VE. (d) Raycasting condition VE. In both conditions, the current target to select is orange, blue shows the currently highlighted target, green is a successfully selected previous target, and red represents a previous miss. The image in the figures appears slightly shifted since we took the pictures from the left-eye camera of Unity. The participants perceived the targets as a circular arrangement.

Upon arrival, participants were asked to fill out a consent form and a demographic questionnaire. Then, the experimenter explained the user study to the participants, going over the virtual hand and raycasting interactions methods and how to select targets with them. While we did not fix participants' head position and orientation, we reset the position of the circular target arrangement after each round of trials, i.e., after 10 target selections. Doing so allowed the participants to face the circle of targets straight on and centered, to avoid potential visual misalignment issues. Participants were also instructed not to change their looking direction while executing the task, and their head movements were monitored by the experimenter to verify that the followed this instruction.

Similar to previous work [7, 8], participants were positioned at the center of an empty room in the VE, and performed the ISO 9241-400:2015 [29] multidirectional selection task with 11 targets. Targets were positioned in a circular arrangement at equal distances from each other as shown in Figure 1(c) and 1(d). The default color of the targets was grey, except for the current target, which was orange. When the cursor was inside of or touched any target sphere, that sphere was highlighted in blue, and the cursor disappeared to avoid providing a potential depth cue to the participant.

Participants were instructed to "select" the targets using their dominant hand with two interaction techniques: virtual hand and raycasting. In the *virtual hand* condition, the participants were asked to reach out with the hand holding the controller to the targets (Figure 1(a)) while not rotating their torso (and thus their head). We placed a 0.5 cm diameter sphere on top of the virtual controller as the cursor. Before starting the experiment, we ensured that each participant could easily reach to the targets at 65 cm [2]. In the *raycasting* condition, participants were instructed to keep their hands around their shoulders so that they could easily select the targets close to them (Figure 1(b)). We rendered a virtual line from the controller center in the participants' pointing position, with the cursor being placed at the tip of the ray, i.e., the point where it intersected geometry. To select a target, participants pressed the "space" bar on the keyboard in front of them with their non-dominant hand. This selection method was chosen to eliminate the "Heisenberg Effect" [13]. If the cursor was inside the target when the user selected it, we recorded a "hit" and turned the color of the sphere green. If not, we recorded a "miss", played an error sound [12], and changed the target color to red.

In our user study, we evaluated three different 3_{VAC} **VAC conditions** in peri-personal space: constant VAC, no VAC, and varying VAC. In the **No VAC** condition, the targets were at the focal plane of the HTV Vive Pro, i.e., 65 cm from the viewer [2], as shown in Fig. 2(b). This 65 cm distance corresponded to 1.54 diopters as the depth distance (100/65 = 1.54). In the **Constant VAC** condition, we used a depth distance of 1.08 diopters to match the difference used in previous work [6, 7]. Thus, the next target depth had to be at (1.08 + 1.54 =) 2.62 diopters, which is equal to $\approx$ 38.2 cm (100/2.62) (Fig. 2(a)). In the **Varying VAC** condition, the spheres were placed in sequence at alternating depths (38.2 and 65 cm) (Fig. 2(c,d)). In other words, if the first target appeared at 65 cm, the next target appeared at 38.2 cm, the third at 65 cm, and so on. As a result of these alternating depths, half of the targets were close to the participants while the other half were far away (see Fig. 2(e,f)).

Previous work on investigating the effect of VAC in stereo displays positioned targets 55 cm away from the participant to analyze lateral movements, i.e., left-to-right and right-to-left, to select targets [3, 6]. They used three different target sizes, 1.5, 2.5, and 3.5 cm and one target distance, 30 cm. Other previous work [7] investigating the effect of VAC on distal pointing also used the exact same target distances and target sizes. Moreover, they varied target distances $\pm$5 cm to widen the ID range and added two more target distances. Still, to analyze their results, they converted Euclidian target size and target distances to angular units, which led to 9 unique angular IDs. Their goal was to obtain the same (visually perceived) size and distance for all targets regardless of their visual depth. Thus, the average distance and size of the target, as observed by the par-

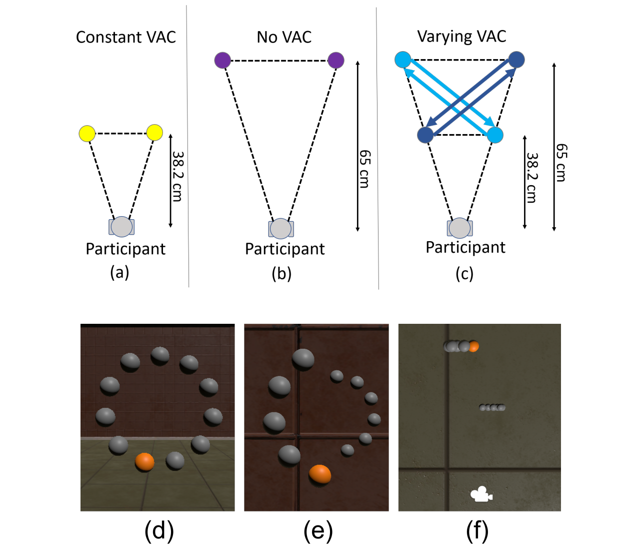


Figure 2: VAC conditions: (a) Constant VAC, where the targets in Fig. 1 appeared at 38.2 cm and are represented with yellow spheres; (b) No VAC condition, where the targets appeared at 65 cm and are represented with purple spheres; and (c) varying VAC condition, where the participant had to select alternating targets at 38.2 cm and 65 cm. For each trial, we randomized the position of the change of depth within the circle of targets. (d) Participants' (perspective) view of targets in the varying VAC condition. The image was taken inside the virtual environment. (e) Isometric view of the varying VAC condition. This image is only used as an illustration of the varying VAC condition, but was never shown to participants. Bigger targets are at the target depth of 65 cm, and smaller targets are at 38.2 cm. (f) Top-down view for the varying VAC condition.

ticipant, was the same regardless of whether there was a constant VAC, constant VAC, or varying VAC. In this paper, we use the same method. Using the same angular IDs for raycasting as in previous work [7], we calculated the corresponding target sizes and target distances at 38.2 and 65 cm, see Table 1. This allowed us to show targets with the same perceived size at both 38.2 and 65 cm, while still maintaining comparability to previous work.

Even though we used this method to specify the target sizes and positions for the raycasting and virtual hand interaction techniques, the motion distance needed to reach the next target is different for both methods. Thus, it was not appropriate to apply the same idea to the varying VAC condition with the virtual hand. In the raycasting method, the participants can sweep the cursor by rotating their wrist. An Euclidean distance corresponding to this angular sweep has to be covered in the virtual hand condition. However, when we look at the angular IDs and Euclidean IDs, we can see that they have the same task difficulty for both interaction techniques. Thus, we decided to use the same target sizes and distances.However, for the varying VAC condition, the distance that had to be covered by the virtual hand method was larger due to the diagonal movement. In other words, A_e increased because of the diagonal distance between targets at 38.2 and 65 cm. To also generate the same ID, we adjusted the size of the target spheres correspondingly in the varying VAC condition based on the Euclidean ID. This method guarantees that participants experienced the same ID for all conditions.

In pilot studies, we observed that participants' fatigue increased with the virtual hand condition. To minimize fatigue, previous work on the VAC [6] required participants to rest for 45 seconds between rounds of trials. Similarly, in this experiment, the participants were forced to rest 10 seconds after each round of 10 selections, and they were given extra time as needed. The resting time never exceeded 60 seconds between trials. Between the conditions, the participants were given a short ≈5 minutes break to rest their eyes and hands.

After the experiment, the participants filled out a post-experiment survey, where they expressed their preferences in terms of interaction techniques and VAC conditions. They also reported their mental and physical fatigue. On average, participants completed the experiment in 20-25 minutes.

3.4 Experimental Design

We conducted a two-factor within-subjects user study with three different **VAC** conditions (3_{VAC} = no VAC, constant VAC, varying VAC) and two **interaction techniques** (2_{IM} = virtual hand and raycasting), yielding a $3_{VAC} \times 2_{IM}$ design. As dependent variables, we measured task execution time (seconds), error rate (%), effective throughput (bits/s), SD_x, and ID_e. We counterbalanced the VAC conditions and interaction techniques across participants with a Latin Square. In total, we used 9 unique ID_As, based on three **angular target sizes** (3_{ATD}) and three **angular target distances** (3_{ATS}). Each participant performed $3_{VAC} \times 2_{IM} \times 9_{ID_A} \times$ 11 repetitions = 594 trials.

4 Results

To analyze the results, we used Repeated Measures (RM) ANOVA in SPSS 24. We considered the data to be normal when the Skewness (S) and Kurtosis (K) of the data distribution were within ±1 [24, 36]. Otherwise, we used log-transform before ANOVA. If the data was not normally distributed after the log-transform, we used ART [51] before ANOVA. We used the Bonferroni method for post-hoc analyses and applied Huynh-Feldt correction when $\varepsilon < 0.75$. The graphs shown in the figures show the mean, and the error bars represent the standard deviation of the mean. We first analyzed the results of the separate interaction techniques, then performed the two-way RM ANOVA.

4.1 Detailed Analyses per Interaction Technique

Fig. 3 shows the results for the virtual hand and raycasting interaction techniques in terms of time, error rate, throughput, SD_x, and ID_e.

4.1.1 Virtual Hand

For the virtual hand interaction technique, ID_e was normally distributed (S = 0.23 and K = 0.14). Time (S = 0.07. K = 0.28), throughput (S = -0.42, K = 0.63), and SDx (S = -0.04, K = -0.2) were normally distributed after log-transform. Error Rate was not normally distributed even after log-transform, so we used ART. The results are shown in Table 2 and Fig. 3.

Time: Participants were significantly faster with the constant VAC condition compared to the no VAC and varying VAC conditions.

Error rate: Participants made significantly more errors with the no VAC condition compared to the constant VAC and varying VAC conditions.

Throughput: Participants' throughput performance significantly increased with the constant VAC condition compared to the no VAC and varying VAC conditions.

SD_x: Participants' accuracy significantly decreased with the varying VAC condition compared to the constant VAC and no VAC conditions. Furthermore, participants were more accurate with the constant VAC condition compared to the other two conditions.

ID_e: Participants' precision increased with the no VAC condition compared to the constant VAC and varying VAC conditions.

Table 1: Target sizes and distances used in this work. The two left-most columns show the angular target sizes and distances used in [7]. The two right-most columns show the *ID*. Note that the Euclidian *ID*s used in the no VAC and constant VAC conditions were the same for the current work.

Angular Target Size (°) [7]	Angular Target Distance(°) [7]	Target Distance (cm) at 65 cm depth	Target Size (cm) at 65 cm depth	Target Distance (cm) at 38.2 cm depth	Target Size (cm) at 38.2 cm depth	Euclidian ID	Angular ID (k=1 in Equation 5)
1.45	30.51	22.72	2.27	13.35	1.33	4.39	4.5
2.42	30.51	22.72	3.82	13.35	2.42	3.69	3.8
3.39	30.51	22.72	5.3	13.35	3.11	3.25	3.37
1.49	25.61	21.02	2.52	12.35	1.48	4.14	4.24
2.48	25.61	21.02	4.22	12.35	2.48	3.453	3.55
3.47	25.61	21.02	5.89	12.35	3.46	3.02	3.12
1.42	35.30	25.03	2.14	14.71	1.26	4.6	4.74
2.36	35.30	25.03	3.57	14.71	2.10	3.9	4.02
3.31	35.30	25.03	5.03	14.71	2.96	3.452	3.59

Table 2: Virtual Hand Data Analysis Results

	VAC	ID
Time	$F(2, 34) = 10.58$, $p < 0.001$, $\eta^2 = 0.384$	$F(8, 136) = 42.71$, $p < 0.001$, $\eta^2 = 0.715$
Error rate	$F(2, 34) = 6.62$, $p < 0.01$, $\eta^2 = 0.280$	$F(8, 136) = 2.81$, $p < 0.01$, $\eta^2 = 0.143$
Throughput	$F(1.49, 25.24) = 9.18$, $p < 0.01$, $\eta^2 = 0.351$	$F(8, 136) = 3.44$, $p < 0.001$, $\eta^2 = 0.168$
SD_x	$F(1.38, 24.79) = 73.0$, $p < 0.001$, $\eta^2 = 0.81$	$F(8,136) = 46.32$, $p < 0.001$, $\eta^2 = 0.732$
ID_e	$F(1{,}363, 23.172) = 8.538$, $p < 0.01$, $\eta^2 = 0.334$	$F(8,136) = 54.788$, $p < 0.001$, $\eta^2 = 0.763$

4.1.2 Raycasting

For the raycasting interaction technique, throughput (S = 0.21, K = 0.58), SD_x (S = 0.72, K = 0.68) and ID_e (S = 0.25, K = 0.58) were normally distributed. Time was normally distributed after log-transform (S = 0.38. K = -0.1). Error Rate was not normally distributed even after log-transform, so we used ART. The results are shown in Table 3 and Fig. 3.

Time: Participants were significantly faster with the constant VAC condition compared to (only) the varying VAC condition.

Error rate: Participants made significantly fewer errors with the varying VAC condition compared to the constant VAC and no VAC conditions.

Throughput: Participants' throughput performance significantly increased with the constant VAC condition compared to (only) the varying VAC condition.

SD_x: Participants' accuracy significantly decreased with the no VAC condition compared to the constant VAC and varying VAC conditions. Furthermore, participants were more accurate with the constant VAC condition compared to the other two conditions.

ID_e: We did not observe any significant difference for participants' precision between the conditions.

Table 3: Raycasting Data Analysis Results

	VAC	ID
Time	$F(2, 34) = 3.64$, $p < 0.05$, $\eta^2 = 0.176$	$F(8, 136) = 90.67$, $p < 0.001$, $\eta^2 = 0.842$
Error rate	$F(2, 34) = 23.56$, $p < 0.01$, $\eta^2 = 0.581$	$F(8, 136) = 8.928$, $p < 0.01$, $\eta^2 = 0.344$
Throughput	$F(2, 34) = 3.2$, $p < 0.05$, $\eta^2 = 0.158$	$F(8, 136) = 15.573$, $p < 0.001$, $\eta^2 = 0.478$
SD_x	$F(2, 34) = 45.90$, $p < 0.001$, $\eta^2 = 0.73$	$F(8,136) = 41.74$, $p < 0.001$, $\eta^2 = 0.711$
ID_e	$F(2, 34) = 1.66$, $p = 0.20$, $\eta^2 = 0.089$	$F(8,136) = 63.82$, $p < 0.001$, $\eta^2 = 0.79$

4.2 Two-way RM ANOVA

In the two-way RM ANOVA analysis, throughput (S = 0.41, K = 0.32) and ID_e (S = 0.26, K = 0.34) were normally distributed. Time (S = 0.22. K = -0.01) and SD_x (S = -0.21, K = 0.04) were normally distributed after log-transform. Error Rate was not normally distributed even after log-transform, so we used ART. The main motivation for this study is to compare different VAC conditions, thus the results between interaction techniques are not presented in detail for *brevity* - except if there were notable results. The results are shown in Table 4 and Fig. 4.

Time: Participants were significantly faster with the constant VAC condition compared to the no VAC and varying VAC conditions.

Error rate: Participants made significantly more errors with the no VAC condition compared to the constant VAC and varying VAC conditions.

Throughput: Participants' throughput performance significantly increased with the constant VAC condition compared to the no VAC and varying VAC conditions.

SD_x: Participants' accuracy significantly decreased with the varying VAC conditions compared to the constant VAC and no VAC conditions. Furthermore, participants were more accurate with the constant VAC condition compared to the other two conditions.

ID_e: Participants' precision increased with the no VAC condition compared to the constant VAC and varying VAC conditions.

Table 4: Two-way Interaction Analysis Results

	Interaction method	VAC	ID
Time	$F(1, 17) = 4.188$, $p = 0.058$, $\eta^2 = 0.207$	$F(2, 34) = 7.917$, $p < 0.001$, $\eta^2 = 0.331$	$F(8, 136) = 122.169$, $p < 0.001$, $\eta^2 = 0.884$
Error rate	$F(1, 17) = 0.59$, $p = 0.45$, $\eta^2 = 0.034$	$F(2, 34) = 7.206$, $p < 0.01$, $\eta^2 = 0.298$	$F(8, 136) = 6.113$, $p < 0.01$, $\eta^2 = 0.264$
Throughput	$F(1, 17) = 2.405$, $p = 0.141$, $\eta^2 = 0.131$	$F(2, 34) = 18.038$, $p < 0.001$, $\eta^2 = 0.53$	$F(8, 136) = 6.79$, $p < 0.01$, $\eta^2 = 0.298$
SD_x	$F(1, 17) = 0.342$, $p = 0.567$, $\eta^2 = 0.021$	$F(2, 34) = 159.144$, $p < 0.001$, $\eta^2 = 0.91$	$F(8,136) = 89.26$, $p < 0.001$, $\eta^2 = 0.848$
ID_e	$F(1, 17) = 25.46$, $p < 0.001$, $\eta^2 = 0.614$	$F(2, 34) = 14.184$, $p < 0.001$, $\eta^2 = 0.47$	$F(8, 136) = 108.58$, $p < 0.001$, $\eta^2 = 0.872$

Two-way Interaction Results: We found a statistically significant difference between interaction techniques and VAC conditions in terms of time ($F(2, 32) = 5.51$, $p < 0.01$ $\eta^2 = 0.256$), throughput ($F(2, 32) = 3.45$, $p < 0.05$, $\eta^2 = 0.177$), and SD_x ($F(2, 32) = 14.57$, $p < 0.001$, $\eta^2 = 0.477$) as shown in Fig. 5. According to these results, raycasting was found to be significantly different than the virtual hand in terms of *(i)* faster selection times in the no VAC and varying VAC conditions, *(ii)* higher throughput in the no VAC

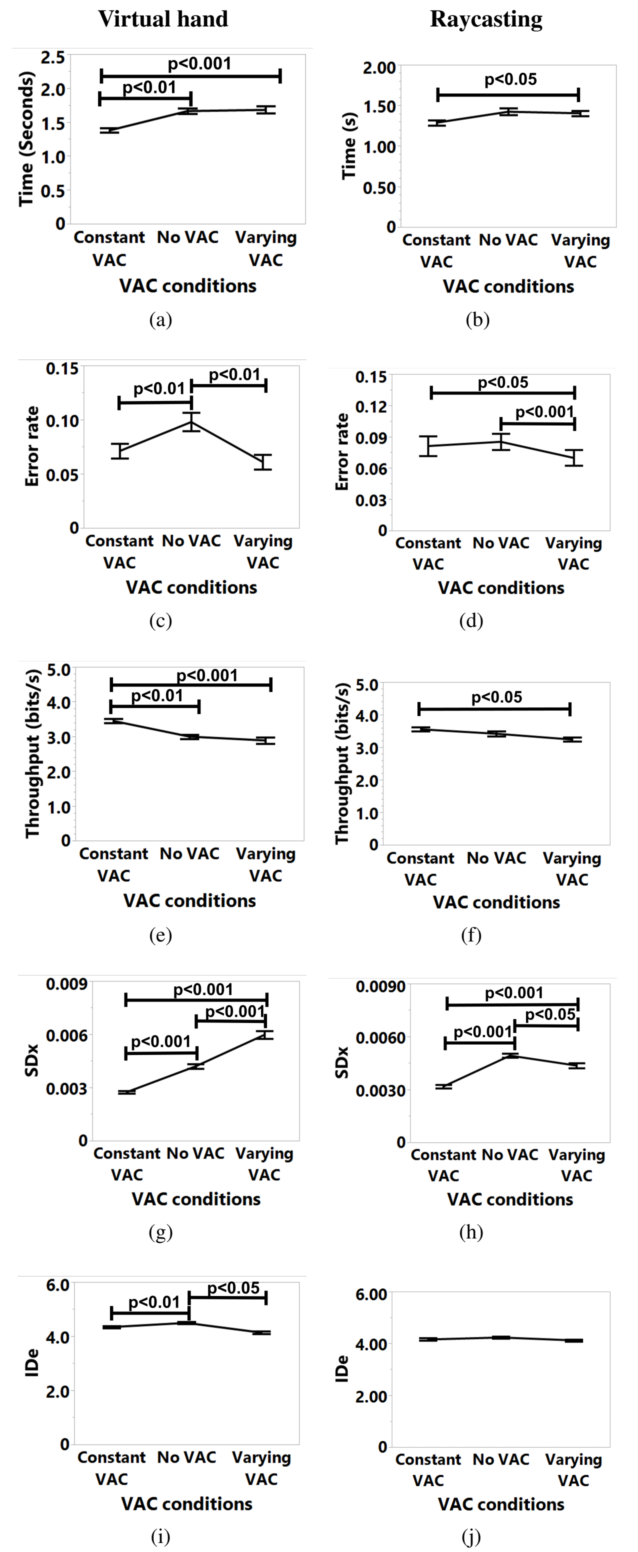


Figure 3: Virtual hand interaction technique results for (a) time, (c) error rate, (e) throughput, (g) standard deviation, and (i) effective index of difficulty. Raycasting interaction technique results for (b) time, (d) error rate, (f) throughput, (h) standard deviation, and (j) effective index of difficulty.

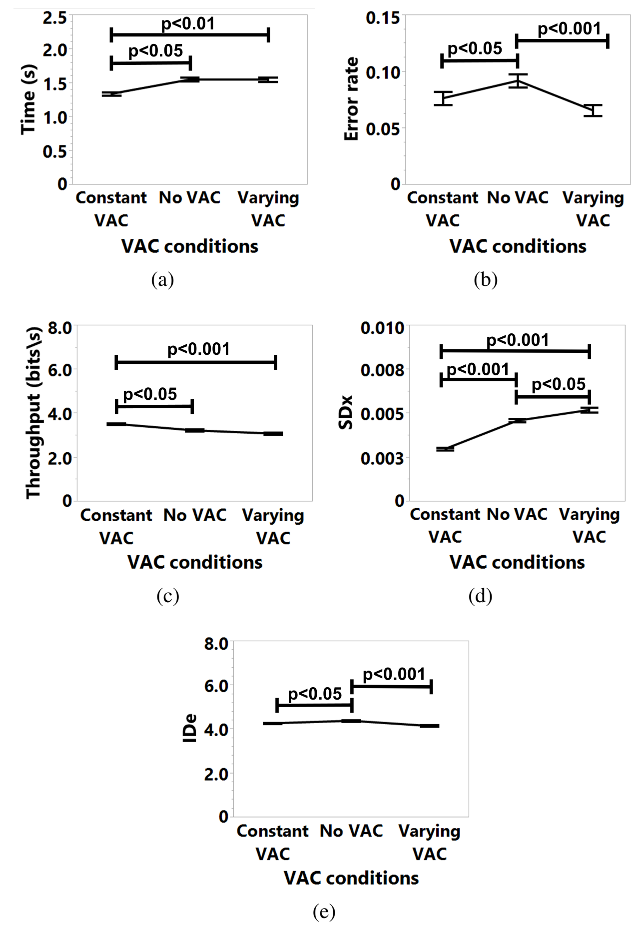


Figure 4: Two-way RM ANOVA results for VAC conditions: (a) Time, (b) error rate, (c) throughput, (d) SD_x, and (e) ID_e.

condition, and *(iii)* better accuracy in the varying VAC condition. On the other hand, the virtual hand exhibited increased accuracy in the constant VAC and no VAC conditions.

4.3 Detailed Varying VAC Condition Analysis

Based on our results, we wanted to analyze the varying VAC condition further and examine how user movement varies while moving between the constant VAC and no VAC target positions. Thus, we ran another RM ANOVA with two **interaction techniques** (2_{IM} = virtual hand and raycasting), two **target positions** (2_{TS} = no VAC and constant VAC), and 9 unique *IDs* (9_{ID}) of the varying VAC condition as shown in Table 5. In this separate analysis, data were normal distributed for throughput (S = 0.42, K = 0.32) and ID_e (S = 0.3, K = 0.1), but for time (S = 0.44. K = 0.19) and SD_x (S = -0.5, K = -0.69) only after log-transform. The error rate was not normally distributed even after log-transform, so we used ART.

Time: Participants were significantly faster when they selected the targets in the constant VAC positions.

Error rate: Participants made significantly more errors when they selected the targets in the constant VAC positions.

Throughput: Participants' throughput significantly increased when they selected targets in the constant VAC positions.

SD_x: Accuracy of the participants significantly decreased when they selected targets in the constant VAC positions.

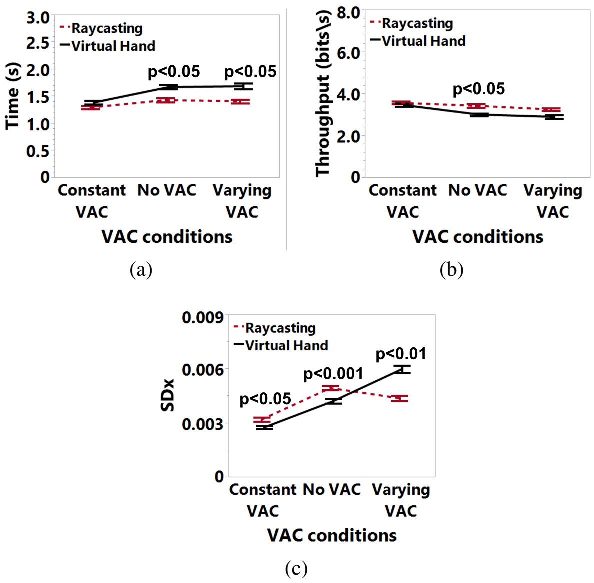


Figure 5: Interaction results for VAC conditions and interaction techniques in terms of (a) time, (b) throughput, and (c) *SDx*.

Table 5: Detailed Varying VAC Condition Results

	Target position	interaction technique	ID
Time	F(1, 17) = 9.619, p<0.01, $\eta^2 = 0.361$	F(1, 17) = 5.163, $p < 0.05$, η^2= 0.233	F(8, 136) = 46.78, $p < 0.001$, $\eta^2 = 0.733$
Error rate	F(1, 17) = 39.8, $p < 0.01$, $\eta^2 = 0.70$	F(1, 17) = 33.182, $p < 0.01$, $\eta^2 = 0.661$	F(8, 136) = 20.077, $p < 0.001$, $\eta^2 = 0.541$
Throughput	F(1, 17) = 32.956, $p < 0.01$, $\eta^2 = 0.66$	F(1, 17) = 4.45, p = 0.136, $\eta^2 = 0.126$	F(8, 136) = 6.170, $p < 0.001$, η^2=0.266
SD_x	F(1, 17) = 136.248, $p < 0.01$, η^2= 0.89	F(2, 34) = 0.12, p = 916, $\eta^2 = 0.01$	F(8, 136) = 18.373, $p < 0.001$, $\eta^2 = 0.525$
ID_e	F(1, 17) = 21.79, $p < 0.01$, η^2= 0.92	F(2, 34) = 0.037, p = 0.849, $\eta^2 = 0.002$	F(8,1 36) = 43.123, $p < 0.001$, $\eta^2 = 0.717$

ID_e: Precision of the participants increased when they selected targets in the constant VAC positions.

Interactions: We found a significant interaction between target position and interaction technique for time (F(1,17) = 8.319, $p < 0.05$, $\eta^2 = 0.329$), error rate (F(1,17) = 24.91, $p < 0.05$, $\eta^2 = 0.594$), and throughput (F(1,17) = 5.182, $p < 0.05$, $\eta^2 = 0.234$). According to these results, participants were faster, made more errors, and had higher throughput performance with the raycasting interaction technique while they selected targets in the constant VAC positions.

4.4 Fitts' Law

We also conducted Fitts' law analysis for both interaction techniques as shown in Fig. 7 and Table 6. Note that the results show a low R^2 = 0.43 value for the no VAC condition for the virtual hand interaction technique.

Table 6: Fitts' Law Analysis Results

	Virtual hand	Raycasting
No VAC	MT = 1.05 + 0.16 × ID R^2 = 0.43	MT = -0.17 + 0.42 × ID R^2 = 0.92
Constant VAC	MT = -0.22 + 0.42 × ID R^2 = 0.97	MT = -0.78 + 0.54 × ID R^2 = 0.97
Varying VAC	MT = -0.23 + 0.5 × ID R^2 = 0.88	MT = -0.37 + 0.47 × ID R^2 = 0.90

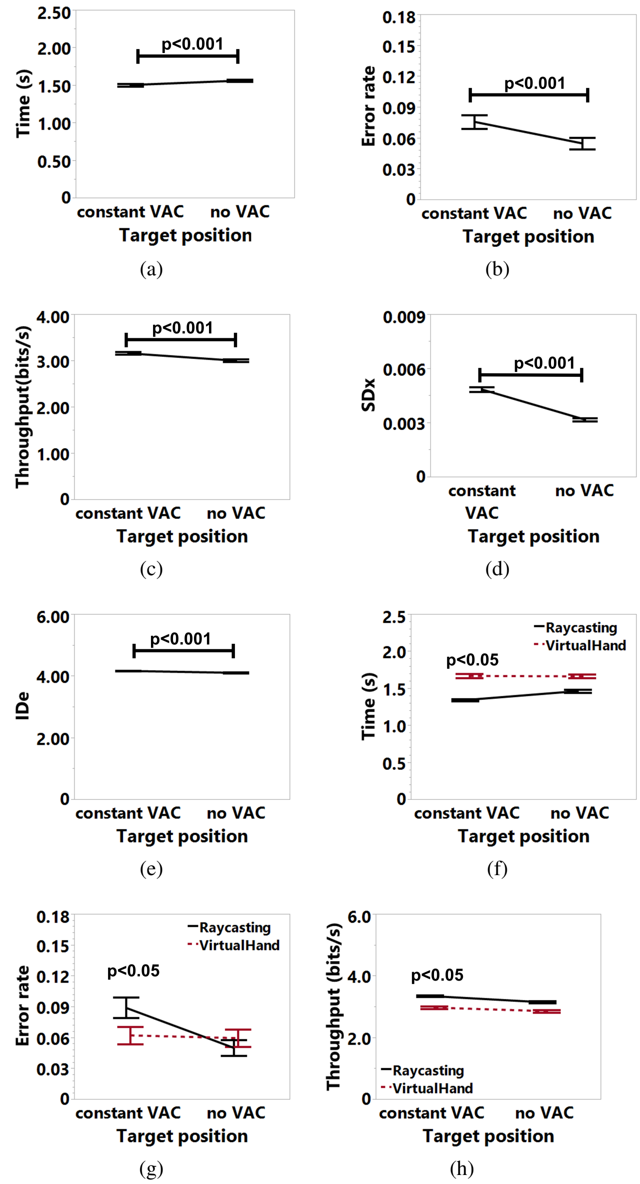
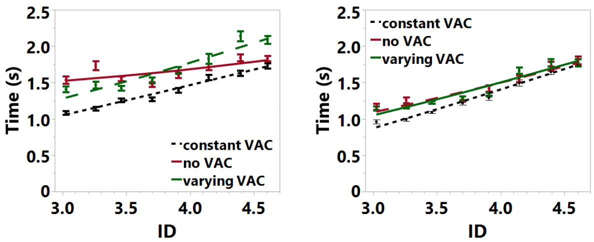


Figure 6: Detailed varying VAC condition results for (a) time, (b) error rate, (c) throughput, (d) standard deviation, and (e) effective index of difficulty. Target position and interaction technique interaction results for (f) time, (g) error rate, and (h) throughput.

4.5 Questionnaire Results

At the end of the experiment, we asked participants a few questions about their preferences between the VAC conditions. We first asked them which VAC condition they preferred. Ten participants preferred the constant VAC condition and supported their choice as "*easier to select the targets and easier to reach out for the targets*." Two participants preferred the no VAC condition and defended their choice as "*less confusing, easier to control, more comfortable, or less headache*." Six participants preferred the varying VAC condition and supported their choice as "*more enjoyable and more realistic*." One participant who preferred the varying VAC condition also commented that "*constant VAC condition was easier to make accurate selections, but it was hard to rotate the cursor/ray through the hoop. Varying VAC felt like a better compromise between the other two.*"



(a) (b)

Figure 7: Fitts' law analysis results for (a) virtual hand and (b) raycasting.

We also asked them if it was easy to select targets with the different VAC conditions (1-I totally disagree, 7-I totally agree) using a 7-point Likert scale. The results showed participants thought it was easier to select targets in the constant VAC condition (mean (M): 6.5, median (Mdn): 7, standard deviation (SD): 0.786), compared to the no VAC condition (M: 4.278, Mdn: 4, SD: 1.565) and the varying VAC condition (M: 4.667, Mdn: 5, SD: 1.782). Finally, we asked them to evaluate their level of physical and mental fatigue after the experiment (1-I feel totally normal 7-I feel completely fatigued). Participants felt neither strong physical fatigue (M: 3.5, Mdn: 3,4, SD: 1.6) nor mental fatigue (M: 2.2, Mdn: 2, SD: 1.1).

5 Discussion

In this paper, we conducted a user study to analyze 3D pointing performance in peri-personal space, i.e., within arm's reach. We asked participants to select targets in an ISO 9241-411:2015 multi-directional task with two frequently used interaction techniques: virtual hand and raycasting. We placed targets at locations where no VAC occurs, where a constant VAC occurs, and also alternated between them, effectively varying the VAC.

5.1 3D Pointing Results

We initially hypothesized that **3D selection with virtual hand and raycasting is negatively affected by the presence of VAC**. Our results partially support our hypothesis, as participants exhibited higher precision (ID_e) with the no VAC condition. However, participants performed better in terms of time, error rate, throughput, and accuracy in the constant VAC condition with both interaction techniques. The questionnaire results support the quantitative analysis, as the majority of the participants preferred the constant VAC condition, and felt targets were easier to select. There are four possible explanations for our results:

Biomechanical Limitations: We hypothesize that both interaction techniques were subject to biomechanical issues which affected our results, but also made results similar. With the **virtual hand** technique, the target placement for the no VAC condition was at 65 cm, i.e., the focal plane of the VR HMD, which is farther away than the constant VAC condition at 38.5 cm. Prior to the experiment, we verified that all participants could reach the targets easily. Yet, the no VAC condition movement involved stretching the hand using the shoulder muscles. This might lead to slower movements compared to simply rotating the elbow, a motion that was predominantly used in the constant VAC and varying VAC conditions for reaching the targets. Further, in contrast to previous work [4, 6, 8] we evaluated diagonal movements, which cross the vertical midline of the body and are more complex than movements only in depth [41, 44]. For the **raycasting technique**, we asked participants to position the controller above and in front of their shoulder to select the targets, as previously shown in Figure 1. Even thought this is not a very common selection pose, we chose this selection pose to allow the participants to select the targets without unusually large rotational movements.

The Fresnel Lens and distortions in the focal plane: The commercial HMD we used for the experiment utilizes Fresnel lenses to render the content, where previous work showed that they distort the focal plane and convert a circle into an elliptical shape [1, 5, 37]. For the targets in the no VAC condition, we believe that the targets' position on the focal plane might have been slightly shifted due to these distortions. Such a shift on the target plane might result in a drop on the participants' performance compared to the other two VAC conditions.

Depth Cues: In our experiment, we intentionally minimized the depth cues that the participants can use in the environment. For instance, we did not render shadows onto the targets and the cursor/ray, targets did not cast shadows into the scene, we used the same shading for all targets regardless of the VAC condition. As perceived cursor size could serve as a depth cue, we also did not render the cursor image when the ray intersected a target. However, other environmental factors (e.g., the length of the ray) could have helped participants to perceive target depth better. This might decrease the effect of the VAC in the constant VAC condition, and reduce the detrimental effects of the VAC in the varying VAC condition.

Task Complexity: Our experiment used a (relatively) more complex task compared to previous work that observed a performance decrease for pointing at distant targets [8]. However, a previous investigation of distal pointing with a multi-directional ISO 9241-400:2015 selection task also showed that the VAC decreases user performance and participants exhibited the worst time and throughput performance in the varying VAC condition [7]. Another potential cause is the specific interaction techniques chosen for our current study, but previous work had already shown that the VAC increases time and throughput performance of the participants in a comparison between singlefocal and multifocal displays, for both raycasting [6] and virtual hand [8]. Still, we took the following measures to reduce potential confounds, such as using a commercial VR HMD, not allowing participants to move during the trials, and resetting the target placement frequently (i.e., after each round of 10 trials) to minimize the perceived depth and position changes for the varying VAC condition. When we analyzed the varying VAC condition in detail, we also observed that results were also similar to previous studies [4, 8], where the participants were slower, made more errors, and their throughput performance, accuracy, and precision increased when they selected targets at 65 cm. However, these additional analyses did not provide additional insights on our current results.

We deliberately chose the 65 cm and 38.2 cm target depth distances in this experiment to match the depth change of 1.08 diopters used in previous work [3, 6–8]. In Barrera Machuca and Stuerzlinger's study [3], the authors showed that the VAC exists in stereo displays, and Batmaz et al.'s follow-up work [8] showed that the VAC cannot be eliminated through environmental cues in AR HMDs. In later work, Batmaz et al. [6] showed that using multifocal displays increases the motor performance of the participants. Then, an investigation of the effect of VAC on distal pointing revealed that the VAC also affects selection performance of distal targets negatively [7]. In all four studies, the authors revealed the detrimental effects of VAC on user performance with a similar depth plane distance, 1.08 diopters.

Our results for pointing in peri-personal space also contradict the results of previous studies with stereo displays. For instance, Teather et al. [48] used a similar experimental design and showed participants targets at different depth distances and asked them to select the targets with raycasting. However, the authors varied the object size so that participants perceived the same target size regardless of the depth distance. Teather et al. [48] also showed that time and throughput of the participants do not change at different depth distances when the perceived target size does not vary.

5.2 Interaction Techniques

In this study, we asked participants to select targets with the two most frequently used interaction techniques, virtual hand and raycasting [32]. However, we did not aim to explore the difference between these interaction techniques, given that it had been previously studied several times, and our results match the outcomes of previous work [9, 18], where participants were reported to be more precise with the virtual hand. Since the objects were closer to the user with the virtual hand, one might conclude that it was easier to select targets. Our Fitts' law analysis for the no VAC condition also supports this claim.

Our main motivation for including two different interaction techniques was to understand how different interaction techniques are affected under different VAC conditions in VR HMDs. The results showed that participants were slower with the virtual hand while selecting targets in the no VAC and varying VAC conditions. Similarly, their throughput performance significantly decreased in the no VAC condition with the virtual hand. Further, we observed higher accuracy with raycasting in the varying VAC condition, while participants were more accurate with the virtual hand in the constant VAC and no VAC conditions. We believe that the reason behind this outcome is that raycasting is more prone to rotational jitter [11], which can negatively affect the user performance in terms of accuracy.

5.3 Other Considerations

As in the other above-mentioned experiments, we did not alter the stimuli in the system except for the changes necessary to create the VAC. In all conditions, target spheres, color, error sounds, and other environmental stimuli were thus the same. We only varied the target position and target size in Euclidean space to match the angular ID across all interaction techniques and VAC conditions. At the beginning of the experiment, we did not inform participants about the properties of the varying VAC condition to eliminate any potential bias towards this condition. We also used a counterbalanced experimental design and did not reveal the names of the experimental conditions to the participants. Thus, we believe that participants were not aware of the precise nature of our experiment. Also, none of the participants reported strong physical or mental fatigue after the study, which makes it unlikely that fatigue for a specific condition might have affected the outcome of the experiment.

From a statistical point of view, one could argue that our results can always be shown to be false; a big enough sample size could almost always lead to statistically significant differences [38]. However, the study did not feature an unusually large number of participants nor a large number of pointing motions (18 participants, 10692 pointing trials). In comparison, Batmaz et al. [6] used 24 participants and 19008 trials. Similarly, in Barrera Machuca and Stuerzlinger's study [3], there were 12 participants and 9504 trials. Overall, the number of collected data thus have a size comparable with previous work. Furthermore, our results exhibit a high effect size $\eta^2 > 0.14$, which indicates a strong effect. In other words, we thus expect that when the experiment is replicated, the probability of observing the same results is high. Overall, our results contribute to the growing body of work on the effect of the VAC on 3D pointing and show that interactions in peri-personal space might be affected by more issues than 'just' the VAC. Even though the effect of the VAC can be observed in a simple task in peri-personal space [6, 8], we suggest that designers should test user performance at various depth distances in a given system and, then, if needed, adapt their system design accordingly.

5.4 Limitations

In this experiment, we only used the HTC Vive Pro and collected data with an ISO 9241-400:2015 task. Our experiment should be replicated with other VR HMDs, such as the Oculus Quest 2. Even though we used the same HMD as previous work on the VAC, such as Batmaz et al. [7], other stereo displays do not necessarily use the same optical design and lens system, i.e., the same focal plane, which affects where the VAC occurs. The Oculus Quest 2, e.g., has a focal plane at 1.3 m and it is thus infeasible for participants to directly reach targets at 1.3 m with their virtual hand.

In this work, we only used two different target depths, i.e., 65 cm and 38.2 cm. Our motivation was to induce the same focal change in terms of diopters, as in previous work, i.e., 1.08 diopters (100/38.2 - 100/65 = 2.62 - 1.54 = 1.08). For instance, Batmaz et al. [7] used depth distances of 75 and 400 cm to elicit the VAC (100/75 - 100/400 = 1.33 - 0.25 = 1.08 diopters). To simplify the experimental design while still enabling comparability, we also focused only on a single change of diopters, like previous work as Batmaz et al. [6]. We still acknowledge the limitation of having used only 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.

Even though we used a VR HMD in this study, our results should also be replicated in AR systems. Previous work on stereo display cue conflicts showed that there is no user performance difference between AR and VR HMDs in terms of time, error rate, and throughput [8]. Still, different HMDs have different display systems. For instance, the Meta AR HMD did not use a lens system but distorted the image on a see-through surface. The HoloLens uses a waveguide to show the virtual content [39]. Our results thus need to be further investigated with AR hardware and other display systems.

In our study, 28% of the participants were male, and 72% were female. This ratio between male and female participants might not be typical of other VR user studies. Thus, we invite researchers to further investigate potential gender differences within VAC studies in the future.

We also used a limited range of *ID*s in this study. Although we selected target sizes and distances to increase the comparability of our results to previous work [7], we suggest extending our experimental design to a larger *ID* range.

6 Conclusion & Future Work

In this paper, we induced the vergence-accommodation conflict (VAC) for pointing movements in a virtual reality (VR) Head Mounted Display (HMD) in peri-personal space. We compared the outcomes when there is no VAC, a constant VAC, and a varying VAC (i.e., when participants had to alternate between these two conditions) with the two most frequently used interaction techniques, virtual hand and raycasting. Interestingly, the results showed that participants were faster and their throughput increased in the constant VAC condition. We speculate that biomechanical constraints, the design of the Fresnel lens used in VR HMD, and other depth cues had an impact on our results. For practitioners, engineers, and designers, we thus suggest to use an approach that involves first positioning the targets at various depth planes, comparing the participants' performance, and then designing the interaction in peri-personal space accordingly.

In the future, we plan to continue analyzing this unexpected phenomenon and to conduct further studies to understand the effect of VAC in peri-personal space with other interaction techniques, different depth distances, and various VR/AR HMDs. We also want to repeat our experiment with a varifocal display and further analyze the VAC in stereo displays.

References

[1] K. Bang, Y. Jo, M. Chae, and B. Lee. Lenslet VR: thin, flat and wide-fov virtual reality display using fresnel lens and lenslet array. *IEEE Transactions on Visualization & Computer Graphics*, 27(05):2545–2554, may 2021. doi: 10.1109/TVCG.2021.3067758

[2] M. S. Banks. [visionlist] focal distance / accommodation distance htc vive pro. http://visionscience.com/pipermail/visionlist_visionscience.com/2021/005947.html.

[3] M. D. Barrera Machuca and W. Stuerzlinger. The effect of stereo display deficiencies on virtual hand pointing. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, p. 1–14. Association for Computing Machinery, New York, NY, USA, 2019. doi: 10.1145/3290605.3300437
[4] M. D. Barrera Machuca, W. Stuerzlinger, and P. Asente. The effect of spatial ability on immersive 3D drawing. In *Proceedings of the 2019 on Creativity and Cognition*, pp. 173–186. ACM, 2019. doi: 10.1145/3325480.3325489
[5] U. Y. Başak, S. Kazempourradi, E. Ulusoy, and H. Ürey. Wide field-of-view dual-focal-plane augmented reality display. In *Advances in Display Technologies IX*, vol. 10942, pp. 62–68. SPIE, 2019.
[6] A. U. Batmaz, M. D. Barrera Machuca, J. Sun, and W. Stuerzlinger. The effect of the vergence-accommodation conflict on virtual hand pointing in immersive displays. In *CHI Conference on Human Factors in Computing Systems*, CHI '22. Association for Computing Machinery, New York, NY, USA, 2022.
[7] A. U. Batmaz, M. M. Hudhud, M. D. Barrera Machuca, and W. Stuerzlinger. Effect of stereo deficiencies on virtual distal pointing. In *28th Symposium on Virtual Reality Software and Technology*, VRST '22. Association for Computing Machinery, New York, NY, USA, 2022.
[8] A. U. Batmaz, M. D. B. Machuca, D. M. Pham, and W. Stuerzlinger. 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)*, pp. 585–592, 2019. doi: 10.1109/VR.2019.8797975
[9] A. U. Batmaz, A. Mutasim K., and W. Stuerzlinger. Precision vs. power grip: A comparison of pen grip styles for selection in virtual reality. In *Conference on Virtual Reality and 3D User Interfaces, Abstracts and Workshops, Workshop on Novel Input Devices and Interaction Techniques at VR*, NIDIT '20, pp. 23–28. IEEE, Mar 2020. doi: 10.1109/VRW50115.2020.00012
[10] A. U. Batmaz, M. Rajabi Seraji, J. Kneifel, and W. Stuerzlinger. No jitter please: Effects of rotational and positional jitter on 3D mid-air interaction. In *Future Technologies Conference*, vol. AISC 1289 of *FTC '20*. Springer, Nov 2020. doi: 10.1007/978-3-030-63089-8_52
[11] A. U. Batmaz and W. Stuerzlinger. The effect of rotational jitter on 3D pointing tasks. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*, CHI EA '19. ACM, May 2019.
[12] A. U. Batmaz and W. Stuerzlinger. 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, pp. 85–94, Mar 2021. doi: 10.1109/VR50410.2021.00029
[13] D. Bowman, C. Wingrave, J. Campbell, and V. Q. Ly. Using pinch gloves (tm) for both natural and abstract interaction techniques in virtual environments. In *HCI International*, pp. 629–633, 2001.
[14] J. M. Brown and N. Weisstein. A spatial frequency effect on perceived depth. *Perception & Psychophysics*, 44(2):157–166, Mar 1988. doi: 10.3758/BF03208708
[15] Y. Cha and R. Myung. Extended Fitts' law for 3D pointing tasks using 3D target arrangements. *International Journal of Industrial Ergonomics*, 43(4):350–355, Jul 2013. doi: 10.1016/j.ergon.2013.05.005
[16] L. D. Clark, A. B. Bhagat, and S. L. Riggs. Extending Fitts' law in three-dimensional virtual environments with current low-cost virtual reality technology. *International Journal of Human-Computer Studies*, 139:102413, 2020. doi: 10.1016/j.ijhcs.2020.102413
[17] F. Daniel and Z. Kapoula. Induced vergence-accommodation conflict reduces cognitive performance in the Stroop test. *Scientific Reports*, 9(1):1–13, 2019. doi: 10.1038/s41598-018-37778-y
[18] P. Duc-Minh and W. Stuerzlinger. 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, Nov 2019. doi: 10.1145/3359996.3364264
[19] F. H. Durgin, D. R. Proffitt, T. J. Olson, and K. S. Reinke. Comparing depth from motion with depth from binocular disparity. *Journal of Experimental Psychology: Human Perception and Performance*, 21(3):679–699, 1995. doi: 10.1037/0096-1523.21.3.679
[20] F. H. Durgin, D. R. Proffitt, T. J. Olson, and K. S. Reinke. Comparing depth from motion with depth from binocular disparity. *Journal of Experimental Psychology: Human Perception and Performance*, 21(3):679–699, 1995. doi: 10.1037/0096-1523.21.3.679
[21] G. N. Dutton, A. Saaed, B. Fahad, R. Fraser, G. McDaid, J. McDade, A. Mackintosh, T. Rane, and K. Spowart. 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):27–34, jan 2004. doi: 10.1038/sj.eye.6700541
[22] P. M. Fitts. The information capacity of the human motor system in controlling the amplitude of movement. *Journal of experimental psychology*, 47(6):381, 1954.
[23] T. Fukushima, M. Torii, K. Ukai, J. S. Wolffsohn, and B. Gilmartin. The relationship between CA/C ratio and individual differences in dynamic accommodative responses while viewing stereoscopic images. *Journal of Vision*, 9(13):21–21, 12 2009. doi: 10.1167/9.13.21
[24] J. F. Hair Jr, W. C. Black, B. J. Babin, and R. E. Anderson. Multivariate data analysis, 2014.
[25] D. M. Hoffman, A. R. Girshick, K. Akeley, and M. S. Banks. Vergence-accommodation conflicts hinder visual performance and cause visual fatigue. *Journal of Vision*, 8(3):33.1–30, mar 2008. doi: 10.1167/8.3.33
[26] A. Huckauf, M. H. Urbina, I. Böckelmann, L. Schega, R. Mecke, J. Grubert, F. Doil, and J. Tümler. Perceptual issues in optical-see-through displays. *Proceedings - APGV 2010: Symposium on Applied Perception in Graphics and Visualization*, 1(212):41–48, 2010. doi: 10.1145/1836248.1836255
[27] M. Hudhud Mughrabi, A. K. Mutasim, W. Stuerzlinger, and A. U. Batmaz. My eyes hurt: Effects of jitter in 3D gaze tracking. In *Conference on Virtual Reality and 3D User Interfaces, Abstracts and Workshops, Workshop on Novel Input Devices and Interaction Techniques at VR*, NIDIT '22, Mar 2022.
[28] J. Iskander, M. Hossny, and S. Nahavandi. Using biomechanics to investigate the effect of VR on eye vergence system. *Applied Ergonomics*, 81(August 2018):102883, 2019. doi: 10.1016/j.apergo.2019.102883
[29] ISO. ISO 9241-400:2015 Ergonomics of human-system interaction - Part 411: Evaluation methods for the design of physical input devices, 2015.
[30] I. Janzen, V. K. Rajendran, and K. S. Booth. Modeling the impact of depth on pointing performance. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, pp. 188–199. ACM, 2016. doi: 10.1145/2858036.2858244
[31] R. Kopper, D. A. Bowman, M. G. Silva, and R. P. McMahan. A human motor behavior model for distal pointing tasks. *International journal of human-computer studies*, 68(10):603–615, 2010.
[32] J. J. LaViola Jr, E. Kruijff, R. P. McMahan, D. Bowman, and I. P. Poupyrev. *3D user interfaces: theory and practice*. Addison-Wesley Professional, 2017.
[33] I. S. MacKenzie. Fitts' law as a research and design tool in human-computer interaction. *Human-computer interaction*, 7(1):91–139, 1992.
[34] I. S. MacKenzie and P. Isokoski. Fitts' throughput and the speed-accuracy tradeoff. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '08, p. 1633–1636. Association for Computing Machinery, New York, NY, USA, 2008. doi: 10.1145/1357054.1357308
[35] I. S. MacKenzie and A. Oniszczak. A comparison of three selection techniques for touchpads. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 336–343. ACM Press/Addison-Wesley Publishing Co., 1998.
[36] P. Mallery and D. George. SPSS for windows step by step: a simple guide and reference. *Allyn, Bacon, Boston,*, 2003.
[37] J. Martschinke, J. Martschinke, M. Stamminger, and F. Bauer. Gaze-dependent distortion correction for thick lenses in hmds. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 1848–1851, 2019. doi: 10.1109/VR.2019.8798107
[38] P. E. Meehl. Theory-testing in psychology and physics: A methodological paradox. *Philosophy of science*, 34(2):103–115, 1967.
[39] Microsoft. HoloLens 2, 2022.
[40] A. Murata and H. Iwase. Extending Fitts' law to a three-dimensional

pointing task. *Human Movement Science*, 20(6):791–805, Dec 2001. doi: 10.1016/S0167-9457(01)00058-6

[41] K. Newell C. *The Slow Learner in the Classroom*. C. E. Merrill Books, Columbus, 1962.

[42] R. Patterson and W. L. Martin. Human stereopsis. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 34(6):669–92, Dec 1992. doi: 10.1177/001872089203400603

[43] R. S. Renner, B. M. Velichkovsky, and J. R. Helmert. The perception of egocentric distances in virtual environments - A review. *ACM Computing Surveys*, 46(2):1–40, Nov 2013. doi: 10.1145/2543581.2543590

[44] W. N. Schofield. Do children find movements which cross the body midline difficult? *Quarterly Journal of Experimental Psychology*, 28(4):571–582, 1976. doi: 10.1080/14640747608400584

[45] R. W. Soukoreff and I. S. MacKenzie. Towards a standard for pointing device evaluation, perspectives on 27 years of fitts' law research in hci. *Int. J. Hum.-Comput. Stud.*, 61(6):751–789, Dec. 2004.

[46] J. Sun and W. Stuerzlinger. Extended sliding in virtual reality. In *25th ACM Symposium on Virtual Reality Software and Technology*, VRST '19. Association for Computing Machinery, New York, NY, USA, 2019. doi: 10.1145/3359996.3364251

[47] R. J. Teather, A. Pavlovych, W. Stuerzlinger, and I. S. MacKenzie. Effects of tracking technology, latency, and spatial jitter on object movement. In *3D User Interfaces, 2009. 3DUI 2009. IEEE Symposium on*, pp. 43–50. IEEE, 2009.

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

[49] A. Thomas. Virtual reality (VR) headset unit sales worldwide from 2019 to 2024. https://www.statista.com/statistics/677096/vr-headsets-worldwide/, 2022.

[50] C. Vienne, L. Sorin, L. Blondé, Q. Huynh-Thu, and P. Mamassian. Effect of the accommodation-vergence conflict on vergence eye movements. *Vision Research*, 100:124–133, 2014. doi: 10.1016/j.visres.2014.04.017

[51] J. O. Wobbrock, L. Findlater, D. Gergle, and J. J. Higgins. The aligned rank transform for nonparametric factorial analyses using only ANOVA procedures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pp. 143–146. ACM, New York, NY, USA, 2011. doi: 10.1145/1978942.1978963