Dynamic Vergence–Accommodation Conflict: Effects on 3D Selection Performance When Target Depth Changes

Zahra Rasoulifar, Rumeysa Turkmen, Mayra Donaji Barrera Machuca,
Wolfgang Stuerzlinger, and Anil Ufuk Batmaz

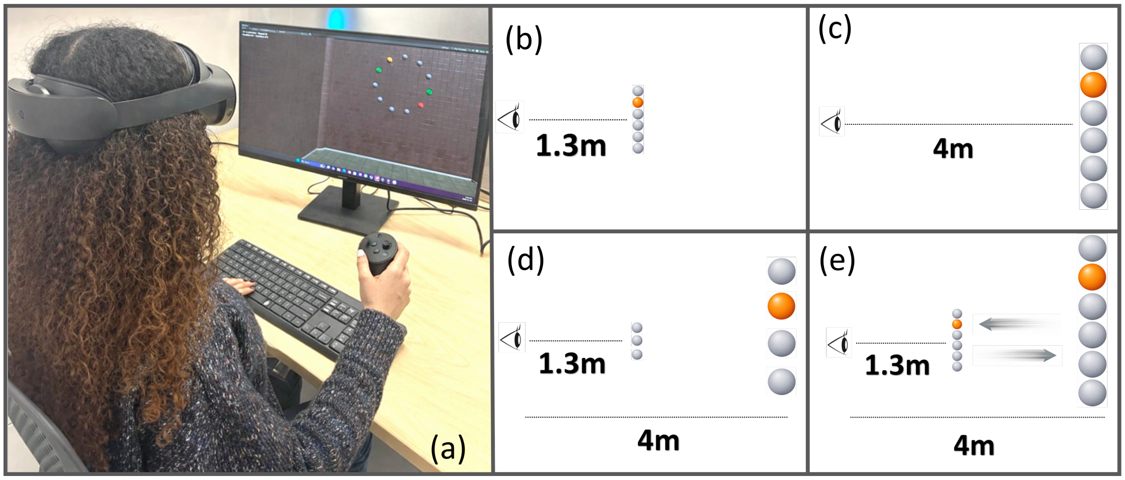


Figure 1: Overview of the experimental setup and Vergence–Accommodation Conflict (VAC) conditions. (a) A participant performing the ISO 9241-411 multidirectional selection task in VR while pointing with a controller and confirming selections using a keyboard. (b) *No VAC*: targets are located at the headset's focal plane (1.3 m), so vergence and accommodation demands match. (c) *Constant VAC*: targets are placed further from the focal plane (4 m), creating a fixed vergence–accommodation conflict during selection. (d) *Varying VAC*: targets appear alternately at near and far depths across selections. (e) *Dynamic VAC*: targets continuously move between depths during acquisition while maintaining a constant angular size, causing the vergence–accommodation conflict to change over time.

Abstract— The effect of the Vergence–Accommodation Conflict (VAC) in state-of-the-art head-mounted displays on 3D interaction performance is well-studied in scenarios where targets remain stationary at a fixed depth. Yet, many VR applications include targets moving continuously toward or away from the user, causing vergence demand to change over time while accommodation remains fixed at the display focal plane. We study this scenario using an ISO 9241-411 multidirectional target selection task with: *No VAC* (targets at the display focal plane), *Constant VAC* (targets further from the focal plane), *Varying VAC* (targets appear at both No VAC and Constant VAC depths across selections), and *Dynamic VAC* (target depth changes continuously during acquisition) conditions. In a within-subjects study with 24 participants, Dynamic VAC produced slower selections, higher error rates, lower accuracy, and reduced throughput compared with the No VAC condition. To account for target movements in depth, we propose an extension of Fitts' law that models the additional performance costs introduced by continuously changing vergence demand. Our results provide guidance for modeling and designing interaction techniques in depth-varying 3D environments.

Index Terms—Vergence-Accommodation Conflict, Virtual Reality, Depth Perception, 3D Interaction, Fitts' law

1 Introduction

Recent progress in extended reality (XR) head-mounted displays (HMDs) has resulted in self-contained, wireless, and lightweight devices that feature low-latency rendering and tracking, a wide field of view (FoV), and up to 4K resolution [4]. These advances have made it possible for industries like healthcare, arts, entertainment, and the military to use XR as a medium for education and to display real-time digital information that users can interact with. For example, previous work has identified the benefits of virtual reality (VR) training in aviation [23], table tennis [36], and surgical skills training [51]. Doctors also use augmented reality (AR) HMDs to help guide them in surgeries [27]. In all these areas, it is important to support high-precision interaction when using XR HMDs.

One challenge for modern XR HMDs is accurately rendering depth cues, enabling users to perceive the placement of an object in three-dimensional (3D) space correctly. Depth perception, the ability to judge spatial distances and objects, relies on a combination of monocular (e.g., size, motion parallax, accommodation, or texture gradient) and binocular cues (e.g., stereopsis or vergence) [24, 49]. Current XR HMDs do not reproduce all depth cues consistently. For example, optical see-through AR displays may additionally produce conflicts when stereoscopic depth and occlusion cues indicate inconsistent spatial relationships [53]. Au et al. [7] found that conflicts between

• *Zahra Rasoulifar is with Concordia University. E-mail: zahra.rasoulifar@mail.concordia.ca.*
• *Rumeysa Turkmen is with Concordia University. E-mail: rumeysa.turkmen@mail.concordia.ca.*
• *Mayra Donaji Barrera Machuca is with the University of Calgary. E-mail: mbarrera@ucalgary.ca.*
• *Wolfgang Stuerzlinger is with Simon Fraser University. E-mail: w.s@sfu.ca.*
• *Anil Ufuk Batmaz is with Concordia University. E-mail: ufuk.batmaz@concordia.ca.*

Manuscript received xx xxx. 202x; accepted xx xxx. 202x. Date of Publication xx xxx. 202x; date of current version xx xxx. 202x. For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org. Digital Object Identifier: xx.xxxx/TVCG.202x.xxxxxxx

disparity-defined depth and physical occlusion caused systematic underestimation of virtual-object distance. In a related study, Kenoyer-Healy et al. [37] found no significant effect of occluder surface type on action-space distance estimates across the tested optical and video see-through displays. Current commercial single-focal stereoscopic displays also suffer from the vergence-accommodation conflict (VAC) [8,32], a mismatch between the vergence (the inward or outward turning of the eyes to align on an object) and accommodation (the adjustment of the eye's lens to optically focus on the object by changing its shape) [2,47], which affects user performance in 3D selection tasks [10,15]. Previous work has shown that the VAC increases cognitive and perceptual strain, which impairs the user's ability to quickly and accurately perceive an object's spatial position [25,32,58], thereby affecting the speed and precision of interaction with the object [11,14,60].

Past work has investigated the effect of the VAC on user performance under different conditions, such as selection tasks occurring in peri-personal space [10,14] or during distal pointing [13]. More recently, Bashar et al. [11] modeled the effect of the VAC across different focal distances, i.e., diopters, showing a decrease in performance with increasing visual depth. One commonality among these studies is that the target depth remained fixed during acquisition, meaning that users selected a stationary target, where the VAC effect for that target remained constant. However, previous work did not consider the case in which targets move along the depth axis, toward or away from the user.

Vision research has found that depth perception of a moving target is more complex than that of a stationary one [45,48], with past work identifying a high degree of variability between individuals [46] and a higher reliance on stereo depth cues than monocular ones [40]. For the selection of a target moving in depth within XR HMDs, the vergence demand changes continuously while accommodation remains fixed at the display's focal plane. Consequently, the magnitude of the VAC also changes continuously during interaction with such a target. This dynamic condition fundamentally alters the visual constraints under which users perform the task, as they must track and interact with a target while the VACs' magnitude changes over time. Such changes can negatively affect how users visually estimate depth, coordinate eye–hand movements during selection, and select targets. As a result, interaction performance measured under fixed-depth conditions cannot fully represent user behavior when interacting with targets that move in depth.

In this paper, we investigate how targets moving along the depth axis affect user performance in VR. Specifically, we examine how a dynamic VAC condition, produced when targets move toward and away from the user, influences target selection performance compared with conditions where targets remain at fixed depths with and without the VAC. Through a controlled, ISO 9241:411 multidirectional selection task, we analyze how a dynamic VAC affects performance, providing insights into how dynamic depth changes influence interaction performance in immersive environments. Our contributions are:

- Introduce a *Dynamic VAC* interaction condition where targets move toward and away from the user, producing a continuously changing vergence–accommodation conflict during interaction.
- Demonstrate that the Dynamic VAC condition leads to slower task completion times, higher error rates, and reduced selection accuracy compared to fixed-depth conditions.
- Propose a performance model that captures the relationship between task difficulty and target depth when targets move along the depth axis.

2 Previous Work

2.1 VAC in Commercial HMDs

Prior work indicates that VAC can affect both visual experience and visual performance. Hoffman et al. showed that the VAC can reduce visual performance and increase visual fatigue [32]. Binocular image imperfections, including spatial distortions, photometric differences between the two eye images, crosstalk, and stereoscopic disparities, can reduce visual comfort when sufficiently large [38]. The VAC has also been connected to changes in visual system behavior (e.g., vergence and accommodation responses) when viewing stereoscopic imagery [29,59] and to measurable costs in demanding cognitive tasks under induced conflict [25].

Commercial VR and AR HMDs like the Quest 3, Varjo-4, and Galaxy XR are effectively single-focal systems due to the use of stereo displays at fixed focal distances as their underlying technology [2, 47]. Stereo displays render two different images, one for each eye, to create the binocular disparity needed to perceive depth. In modern HMDs, the physical focal distance is fixed by the display optics. As a result, accommodation is driven toward the display's focal plane, while vergence follows the depth simulated by the stereo content. This mismatch between accommodative distance and vergence distance is known as the VAC [8,16].

In optical see-through AR, focal mismatches can also happen when users switch between real and virtual information at different depths. Arefin et al. [5] found that larger focal switching distances increased eye fatigue and reduced visual-search performance, while transient focal blur caused additional performance costs. Other work has studied multifocal and varifocal displays as approaches to reducing the VAC. MacKenzie et al. [42] found that depth-filtered multifocal stimuli produced accurate accommodation responses when adjacent focal planes were separated by approximately one diopter or less. Multifocal and varifocal displays have also been shown to improve 3D pointing performance relative to single-focal viewing [12,33]. However, these approaches have not been widely adopted in current commercial HMDs.

2.2 VAC and 3D Interaction Performance

Stereo displays can improve 3D target selection, but performance in 3D remains lower than in comparable 2D pointing tasks [55]. For example, several studies have reported that stereo display limitations and depth-related factors reduce selection performance, including slower movement time and reduced throughput (a measurement for speed-accuracy), for targets that require movement in visual depth [9,10,15].

In optical see-through AR, Au et al. [6] found that focal-distance conflicts between real and virtual objects reduced the accuracy of a visually guided ring-placement task, due to systematic underestimation of virtual-object distance relative to physical objects.

Past work has identified that the VAC affects user performance (speed and accuracy) during 3D pointing and selection. For example, Batmaz et al. examined virtual-hand pointing under different VAC conditions and reported significant effects of the conflict on 3D selection performance, particularly for movements that involved (stationary) targets at various points along the depth axis [12]. In a follow-up study, the researchers showed that stereo deficiencies and target position relative to the focal plane significantly affected raycasting selection performance [13]. Past work also identified that biomechanical and ergonomic factors could increase the detrimental effects of VAC during 3D selection [16].

2.3 Depth-Varying and Multi-Depth Interaction

Depth itself is a contributor to pointing difficulty. Work on 3D pointing reports that movement direction and depth placement can affect movement time [35,44,55] and error rates [35,55], and that movements along the depth axis can behave differently than lateral ones [44,55]. Past work also observed the effect of depth for distal pointing, where distance-related factors can change effective measures and user strategies [35,39].

Other studies used multiple depth planes or targets placed at different depths to study selection when visual depth changes. Examples include a varying VAC condition, where target positions switch between the HMD's focal plane and locations away from it [13,16]. These works have shown that the VAC affects the selection performance of the user. Yet, in all conditions, the position of the selected target remained constant during each movement. In other words, the target was rendered at a fixed depth until selection was completed. This is an important limitation of these works, as they do not consider targets that are (continuously) moving in depth.

For studies that focus on dynamic targets, past work has most often focused on the effects of monocular depth cues (texture, shadow) and target characteristics (alignment, speed, and moving direction) [34]. Zheng et al. [62] proposed a three-dimensional model of pointing-endpoint uncertainty for moving targets in VR that incorporates depth-related effects on visual perception and arm-movement stability. Cashion et al. [20] compared four 3D selection techniques across scenarios with different object densities and motion dynamics. Raycasting selects an object by pointing a virtual ray toward it. SQUAD uses progressive refinement by repeatedly presenting groups of candidate objects in quadrants until the desired object can be selected. Zoom extends Raycasting by magnifying the region containing potential targets, while Expand places candidate objects in a grid for final selection. Their results showed that Raycasting and SQUAD had limitations in dense and dynamic environments, while the modified Zoom and Expand techniques improved performance in several of the tested scenarios. Haan et al. [26] proposed IntenSelect for dynamic scenes, which uses a "dynamic object rating" to bend a selection ray toward a target based on its movement over time. However, despite these advancements in handling the spatial and temporal difficulties of moving targets, these evaluations focused primarily on speed, accuracy, and density, and they did not consider the impact of the VAC on user performance during dynamic selection tasks.

2.4 Fitts' Law and 3D Selection in VR

Fitts' law relates movement time to target distance and size and has been widely used for modeling pointing performance in HCI [28, 41]. One of the most commonly used Fitts' law formulations is Shannon's formulation, as shown in Equation 1.

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

In Equation 1, the logarithmic term indicates the *I*ndex of *D*ifficulty (*ID*). The *ID* is calculated from target distance (A) and target size (W). a and b are derived with linear regression.

In HCI research, performance is commonly summarized using throughput, defined as the ratio of effective index of difficulty to movement time, where the effective index of difficulty is computed from effective distance and effective width to capture the speed–accuracy trade-off [52]. The ISO 9241-411 standard [1] defines a multi-directional selection task for evaluating pointing performance with Equation 2.

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

where effective throughput is calculated via the effective ID, ID_e, defined in Equation 3, divided by the movement time (MT).

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

In Equation 3, the A_e represents the effective movement amplitude and W_e the effective target width. W_e is determined by the standard deviation of the distance between selection points and the target center (SD_x), which also indicates the accuracy in the selection performance [12]. Effective throughput and ISO 9241:411 multidirectional selection tasks have been widely adopted in HCI research [52] and for VR studies investigating VAC [16].

For 3D interaction, different formulations of Fitts' law have been proposed, as interaction techniques can involve different movement characteristics and control parameters [10, 11, 21, 22, 44]. For raycasting and other rotational pointing techniques, angular formulations are often used as the motor action is largely wrist rotation and the relevant geometry is better captured by angular separation and angular target width [39]. These modeling choices become especially important when targets are placed at different depths, since the geometric distance to the target and its appearance change with depth. The angular Fitts' law formulation is shown in Equation 4.

$$MT = a + b \cdot \log_2\left(\frac{\alpha_e}{\omega_e^k} + 1\right) = a + b \cdot \log_2\left(\frac{\alpha_e}{(4.133 \cdot SD_x)^k} + 1\right) \tag{4}$$

As with Fitts' law, angular Fitts' law is based on the effective angular index of difficulty, where α_e represents the angular distance and ω_e the angular target width, i.e., the distribution of the angular selection coordinates, calculated as $\omega_e = 4.133 \cdot SD_x$. The constant k represents the relative weight, which is typically set to 1 [16].

2.5 Controlling Visual Angle and Retinal Size

Fitts' law shows that movement time is modeled as a function of movement amplitude and target width [28, 52]. When comparing performance across depth conditions, targets placed closer to the viewer naturally appear larger than targets positioned farther away. As a result, faster performance at nearer depths may reflect increased visibility rather than changes in VAC. Depth-related differences in 3D pointing performance have also been reported in prior VR studies [55]. To address this, previous VAC studies scaled target size and target distance to maintain a constant visual angle across depth conditions, ensuring that the targets subtended the same retinal size across depth planes [12, 14, 16].

Such control of visual angle is especially important when depth changes continuously during a trial. Without scaling, depth motion would cause targets to visibly grow and shrink, adding a clear size change on top of the intended depth manipulation. By keeping angular target position and retinal target size constant, one can vary vergence demand over time while minimizing additional size-change cues during target selection.

3 Motivation & Research Gap

Prior work shows that the vergence-accommodation conflict (VAC) affects 3D selection performance in VR. Single-focal stereo HMDs impose a fixed accommodative demand while vergence varies with simulated depth [8, 16, 32]. As a result, performance results in these studies showed that the VAC affects movement time, accuracy, and throughput, where targets switch discretely between depth planes [12, 13, 16]. However, existing work assumed that the target depth remains constant during each individual selection. Even in varying VAC conditions, depth changes occurred between selections, not during target acquisition itself.

Selecting a target moving in depth is different from selecting a stationary one, due to the increased demands on the visual system [45, 48]. When a target stays at a fixed depth, vergence demand is constant during the movement. In contrast, when the target moves toward or away from the user, vergence demand changes continuously throughout the acquisition. However, accommodation in current single-focal HMDs remains fixed at the display's focal distance. This means that in a Dynamic VAC condition, where the targets move toward or away from the user, the conflict between vergence and accommodation is not constant but evolves during the selection. The visual system must thus repeatedly adjust eye alignment while the optical focus stays fixed. This ongoing adjustment may introduce additional instability or delay compared to static depth conditions. As a result, the Dynamic VAC condition is not simply a stronger version of the static VAC one. It represents a temporally changing conflict that may affect both the perception of depth and motor control during target selection.

As a result, even without noticeable visual motion, the underlying vergence-accommodation relationship evolves during the movement. It is therefore unclear whether the performance effects observed with stationary targets extend to targets that continuously change depth during acquisition. Based on the above, our work is guided by the following research question (**RQ**): *How does continuously changing target depth during acquisition, which creates a dynamic vergence–accommodation conflict, affect 3D target selection performance in a single-focal VR HMD?* We hypothesized that the *Dynamic VAC* condition would negatively affect selection performance compared to conditions with stationary targets, resulting in longer movement times and higher error rates. The corresponding null hypothesis was that selection performance would not significantly differ across these conditions.

4 METHODOLOGY

4.1 Participants

We conducted a user study with 24 participants (12 female, 12 male), aged between 23 and 38 years ($M = 27.54$, $SD = 3.72$). Twenty-one participants were right-handed, and three were left-handed. All participants reported normal or corrected-to-normal vision. Three participants reported no prior VR experience, six tried VR once, four had tried VR twice, one tried VR four times, and ten reported using VR five or more times. All procedures were approved by Concordia's University research ethics board. Participation was voluntary and uncompensated, and no participants were excluded from the study.

4.2 Apparatus

The experiment was conducted on a PC equipped with an Intel(R) Core(TM) i7-14650HX processor, 32 GB of RAM, and an NVIDIA graphics card. We used a *Meta Quest Pro* HMD for the headset, which has a focal distance of 1.3 m [11]. The virtual environment was developed using *Unity* version 2021.3.24f1 and displayed in the HMD via Meta Quest Link.

4.3 Procedure

The user study was conducted on-site in an indoor laboratory on Concordia University campus. After completing the consent and demographic questionnaire, an experimenter explained the study procedure and assisted participants with putting the HMD on. Participants were seated throughout the experiment and were instructed to hold the VR controller in their dominant hand (Figure 1.(a)). The controller was used to point at targets, while the non-dominant hand rested on a keyboard placed in front of them. Participants confirmed selections by pressing the space bar with their non-dominant hand to reduce the "Heisenberg effect" [18], where mechanical input on the controller can unintentionally disturb pointing accuracy.

We used a variation of the ISO 9241-411 multidirectional selection task [1] and adapted it to consist of a ring of spherical targets. Participants were presented with a circular arrangement of gray spheres, with one sphere highlighted in orange as the current target (Figure 1(a)). Each participant completed four experimental blocks, corresponding to the four VAC conditions (4_{VAC}) with this target arrangement:

• *No VAC*: Targets were placed at the focal plane of the VR HMD, i.e., 1.3 m (Figure 1(b)).

• *Constant VAC*: Targets were placed at the far depth of 4 m (Figure 1(c)).

• *Varying VAC*: Half of the targets were placed on the focal plane and the other half at the far distance (Figure 1(d)).

• *Dynamic VAC*: Targets moved continuously back and forth between the focal plane of the VR HMD and the far distance. The depth changed smoothly over 5 seconds to move from the near depth to the far one and vice versa (Figure 1(e)). We selected this duration based on pilot testing as the shortest transition that kept the depth change subtle, so that visible motion cues were minimized while target depth and vergence demand changed continuously.

Across the conditions, we used the headset focal plane at 1.3 m as the near depth [11] and 4 m as the far depth [13]. Following prior VAC work with raycasting [13, 16], we selected this depth range starting at the focal plane and extending further away, which also supports the comparability of our results with such previous work.

Target selection was performed using a small spherical pointer controlled by the VR controller. The pointer was scaled across depths so that its retinal size remained constant. Although the pointer position was computed using an invisible ray, the ray itself was not rendered to avoid making depth changes of the target visually noticeable, particularly in the Dynamic VAC condition. Selection was determined by computing the distance between the pointer and the target center. When this distance became smaller than a predefined proximity radius around the target, the target changed color from orange to blue, indicating that the pointer was inside the selectable area. Correct selections were indicated by the target turning green, while incorrect selections (i.e., confirmations outside the proximity area) were indicated by the target turning red and accompanied by a short auditory cue.

The ring of targets was anchored to the participant's head position and orientation, ensuring that it remained directly in front of them even if small head movements occurred. This head-anchored configuration reduces the possibility of revealing depth differences through head motion, particularly in the Varying VAC and Dynamic VAC conditions. At the start of the experiment, participants were also asked to minimize head movement during selections to maintain consistent task geometry and reduce unnecessary difficulty.

At the start of the session, participants were informed that they would experience four different conditions, without being told how they differed. Before beginning the main experiment, participants completed a training phase in which they practiced the selection task until they felt comfortable with the interaction and understood the task mechanics. During training, all participants correctly identified the color-coded target states and completed the stereoscopic selection task without reporting difficulty perceiving the target depths. Training data was not included in the analysis. The order of the four experimental conditions was fully counterbalanced across participants to reduce order effects.

During the experiment, we measured task execution time (seconds), error rate (%), angular effective THP (bits/s), and angular SD_x. After completing each block, participants removed the headset and completed questionnaires assessing their experience in that condition, including the System Usability Scale (SUS) [19], NASA TLX [31], and additional custom questions related to visual comfort and perceived difficulty. These custom questions asked whether participants experienced difficulty focusing on the targets, noticed any blur or double vision, or felt eye strain during target selection. At the end of the experiment, participants completed a final questionnaire assessing overall physical and mental fatigue, eye discomfort, motion sickness symptoms, and their perception of target depth changes. The full session, including instructions, training, experimental blocks, questionnaires, and breaks, lasted approximately 40–60 minutes per participant.

Participants were given short breaks between VAC conditions, during which they completed questionnaires related to the preceding condition before continuing the experiment.

4.4 Experimental design

We used a within-subjects design with four VAC conditions (4_{VAC}): *No VAC*, *Constant VAC*, *Varying VAC*, and *Dynamic VAC*. All participants completed four conditions with the above-described variation of the ISO 9241:411 multidirectional selection task [1] with 11 targets. To vary the angular index of difficulty (ID_α, henceforth ID for brevity), we used 3 angular target sizes ω_e (0.66°, 1.10°, and 1.54°, corresponding to W : 1.5 cm, 2.5 cm, and 3.5 cm metric size at the 1.3 m focal plane), and 3 angular target distances α_e (11.0°, 13.2°, and 15.4°, corresponding to L : 25 cm, 30 cm, and 35 cm metric units at the 1.3 m focal plane, and combined them, yielding 9_{ID_A}. These physical target sizes and spacings follow parameter ranges used in a prior VR pointing study investigating stereo deficiencies and vergence–accommodation conflict [13]. Within each VAC condition and repetition, the nine ID configurations were presented in randomized order. Overall, each participant completed $4_{VAC} \times 3$ repetitions $\times 9_{ID} \times 11$ trials $= 1188$ selections.

5 RESULTS

We analyzed the data using repeated-measures (RM) ANOVA in SPSS 29, with the VAC condition (4 levels: No VAC, Constant VAC, Varying VAC, Dynamic VAC) as the within-subject factor. Data were considered normally distributed when Skewness and Kurtosis values were within ± 1 [30, 43]. When this assumption was violated (even after log transformation), we applied the Aligned Rank Transform (ART) [61] before analysis. Post hoc comparisons were conducted using Bonferroni correction. Results are reported as means with standard error of the mean (SEM) in the figures. Non-parametric data were analyzed

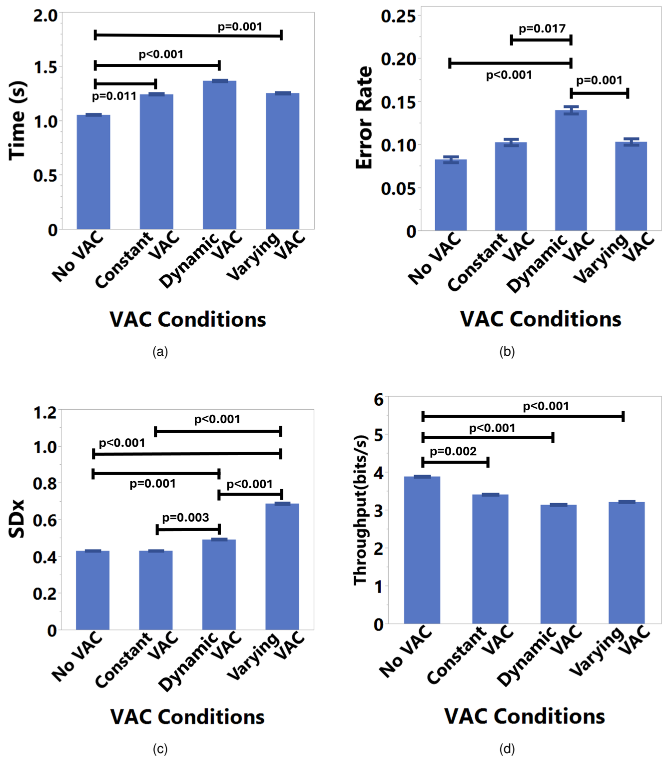


Figure 2: RM ANOVA results for (a) time, (b) error rate, (c) SD_x, and (d) throughput

with a Friedman's test and a Wilcoxon Signed Rank test for pairwise comparisons.

5.1 Performance Results

Movement Time (MT). Movement time was significantly different among VAC conditions (Table 1). As shown in Fig. 2a, the *Dynamic VAC* condition had the largest movement times overall. The participant were significantly faster with the *No VAC* condition compared to *Constant VAC* ($p = 0.011$), *Varying VAC* ($p = 0.001$), and *Dynamic VAC* ($p < 0.001$).

Error Rate. Error rate was significantly different across VAC conditions (Table 1). As shown in Fig. 2b, post hoc comparisons indicated that *Dynamic VAC* resulted in significantly more errors compared to the *No VAC* ($p < 0.001$), *Varying VAC* ($p = 0.001$), and *Constant VAC* ($p = 0.017$) conditions.

SD_x. Angular endpoint variability differed significantly across VAC conditions (Table 1). As illustrated in Fig. 2c, post hoc comparisons showed that *Varying VAC* had significantly higher SD_x than *No VAC* ($p < 0.001$), *Constant VAC* ($p < 0.001$), and *Dynamic VAC* ($p < 0.001$). In addition, *Dynamic VAC* also had significantly higher SD_x than the *No VAC* ($p = 0.001$) and *Constant VAC* ($p = 0.003$) conditions.

Throughput. Throughput differed significantly across VAC conditions (Table 1). As shown in Fig. 2d, the *No VAC* condition exhibited overall the highest throughput. Post hoc comparisons indicated that throughput in the *No VAC* condition was significantly higher than in the *Constant VAC* ($p = 0.002$), *Dynamic VAC* ($p < 0.001$), and *Varying VAC* ($p < 0.001$) conditions.

5.2 Fitts' Law Model Fit

Figure 3 shows the relationship between movement time and ID. When the data from the whole study are analyzed together, linear regression results yield $R^2 = 0.74$ with $a = -3.00$, $b = 1.17$ (Fig. 3a). A separate analysis for each VAC condition also confirmed that movement time increases as ID increases (Fig. 3b). Linear regression results indicated ($a = -0.14$, $b = 0.34$, $R^2 = 0.83$) for the *No VAC* condition, ($a = 0.01$, $b = 0.43$, $R^2 = 0.76$) for *Constant VAC*, ($a = 0.04$, $b = 0.44$, $R^2 = 0.73$) for *Varying VAC*, and ($a = 0.438$, $b = 0.501$, $R^2 = 0.63$) for the *Dynamic VAC* condition.

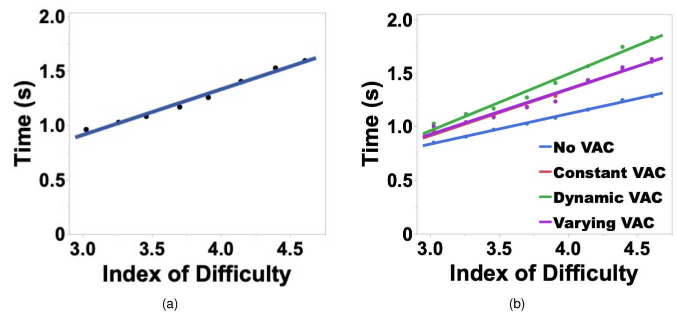


Figure 3: Fitts' law analysis for (a) the whole user study, and (b) each VAC condition.

The relatively low coefficient of determination ($R^2 = 0.63$) observed in the *Dynamic VAC* condition indicates that the standard Fitts' law model does not adequately capture performance in this scenario. Unlike the other experimental conditions, in *Dynamic VAC*, target depth does not remain fixed during acquisition, but changes continuously throughout the movement. Thus, movement time in the Dynamic condition depends not only on ID but also on how the target depth evolves relative to the focal plane during the selection. We argue that a 2-dimensional (ID vs time) linear regression result does not adequately capture the effect of depth. For this reason, we further analyze the Dynamic VAC condition using a depth-dependent representation in the next section.

5.3 Depth-dependent relationship between ID and selection time

To further examine how movement time changed jointly with task difficulty and target depth in the Dynamic VAC condition, we first visualized the empirical data as a smoothed movement-time surface over ID and depth (Figure 4). The empirical surface suggests that movement time generally increases with both variables, while also indicating that the effect of ID is not uniform across depth.

To formalize this pattern, we modeled movement time as a depth-dependent function of ID. Since the headset focal plane was fixed at 1.3 m, depth was expressed relative to the focal plane as

$$\Delta = \max(D - 1.3, 0), \tag{5}$$

where D is the target depth in meters.

We then evaluated five candidate formulations based on prior work and extensions of Fitts' law: the classical Fitts' law formulation [28,41], a linear depth–interaction model, a diopter-based VAC model [11], a power depth model, and a quadratic depth–curvature model designed to capture how the relationship between task difficulty and movement time changes as depth increases.

Table 2 summarizes the comparison of these models. To compare models with different functional forms and numbers of parameters, we used the Akaike Information Criterion (AIC) [3] and the Bayesian Information Criterion (BIC) [50]. Lower AIC and BIC values indicate a better model fit, while also penalizing unnecessary model complexity. These criteria are useful for comparing nonlinear models, where R^2 alone may not provide a reliable basis [54]. Among them, the quadratic depth–curvature formulation provided the best overall fit to the Dynamic VAC data. It achieved the lowest AIC and BIC values, and the highest explained variance ($R^2 = 0.840$, AIC $= -981.90$, BIC $= -964.84$).

The fitted model is therefore:

Parameters	VAC Condition	ID	VAC Condition × ID
Time (s)	**F(3, 69) = 11.559, p<0.001, η^2 = 0.334**	**F(7, 161) = 9.476 p<0.001, η^2 = 0.292**	**F(21, 483) = 5.572 p<0.001, η^2 = 0.195**
Error Rate	**F(3, 69) = 8.890, p<0.001, η^2 = 0.279**	**F(7, 161) = 29.684, p<0.001, η^2 = 0.563**	**F(21, 483) = 54.724 p<0.001, η^2 = 0.704**
SD_x	**F(3, 69) = 2.102, p<0.001, η^2 = 0.819**	**F(7, 161) = 82.196 p<0.001, η^2 = 0.781**	**F(21, 483) = 2.905 p<0.001, η^2 = 0.112**
THP (bits/s)	**F(3, 69) = 16.007, p<0.001, η^2 = 0.410**	**F(7, 161) = 24.947 p<0.001, η^2 = 0.520**	F(21, 483) = 1.243 p=0.210, η^2 = 0.051

Table 1: ANOVA results for VAC condition, index of difficulty (ID), and their interaction across performance measures. Significant effects ($p < 0.05$) are shown in bold.

Table 2: Comparison of candidate models for the Dynamic VAC condition. Lower AIC [3] and BIC [50] values indicate better model fit. Note that for the nonlinear models in the last two rows, R^2 should *not* be used as a measure for comparison [54] and these entries are thus shown in grey.

Model	Formula	Coefficients	R^2	AIC	BIC
1. Classical Fitts' Law [28, 41]	$\mathrm{MT} = a + b \cdot ID$	$a = 0.438$, $b = 0.501$	0.63	−792.4	−784.9
2. Diopter-Based (VAC) [11]	$\mathrm{MT} = a + b \cdot ID + c \cdot \lvert ViD \rvert$	$a = 0.394$, $b = 0.476$, $c = 0.853$	0.76	−893.7	−882.0
3. Linear Depth–Interaction	$\mathrm{MT} = a + b \cdot ID + c \cdot \Delta + d \cdot ID \cdot \Delta$	$a = 0.381$, $b = 0.503$, $c = 0.712$, $d = -0.321$	0.79	−921.4	−906.4
4. Power Depth Model	$\mathrm{MT} = a + b \cdot ID + c \cdot (\Delta + 1)^f$	$a = 0.368$, $b = 0.498$, $c = 0.591$, $f = 1.47$	0.81	−948.3	−933.3
5. Depth-Curvature (Proposed)	$\mathrm{MT} = a + b \cdot ID + c \cdot \Delta + d \cdot ID \cdot \Delta + e \cdot \Delta \cdot ID^2$	$a = -0.3617$, $b = 0.4324$, $c = 0.7110$, $d = -0.3933$, $e = 0.0560$	0.84	**−981.9**	**−964.8**

$\Delta = \max(D - 1.3,\ 0)$; focal plane fixed at 1.3 m. The proposed model is highlighted.

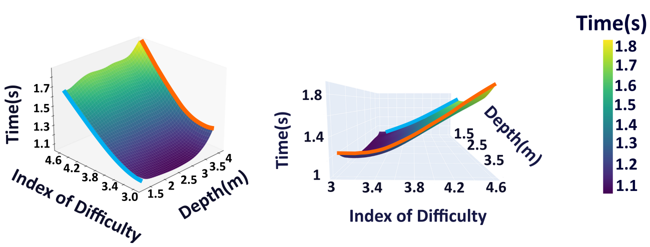


Figure 4: Smoothed empirical movement time surface for the Dynamic VAC condition as a function of index of difficulty and target depth. Color indicates movement time in seconds. The same surface is shown from two viewing angles (front and side) to illustrate the relationship between task difficulty and depth better.

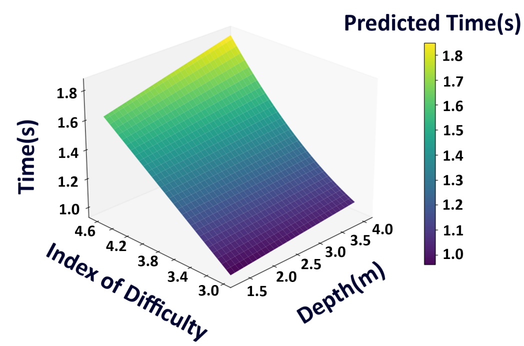


Figure 5: Movement time predicted by the fitted quadratic depth–curvature model as a function of index of difficulty and target depth in the Dynamic VAC condition. Depth is expressed relative to the headset focal plane (1.3 m). The model predicts movement time increases with both task difficulty and depth, while the ID–time relationship becomes increasingly non-linear at farther depths.

$$MT(I, D) = -0.361 + 0.432ID + 0.711\Delta - 0.393ID\Delta + 0.056\Delta ID^2 \tag{6}$$

where MT is the predicted movement time in seconds and ID is the angular index of difficulty. This formulation extends the linear Fitts-style relationship by allowing the curvature of the ID–time function to increase with distance from the focal plane.

Figure 5 shows the movement-time surface generated from the fitted model. The predicted surface captures the same overall trend seen in the empirical data: movement time increases with both depth and ID, while the shape of the ID–time relationship becomes increasingly non-linear at greater depths.

To make this easier to interpret, Figure 6 shows model-based 2D slices at three representative depths: 1.3 m, 2.5 m, and 4.0 m. At the focal plane ($D = 1.3$ m), $\Delta = 0$, and the fitted model reduces to:

$$T(I) = -0.3617 + 0.4324I. \tag{7}$$

At 2.5 m, $\Delta = 1.2$, and the fitted model becomes:

$$T(I) = 0.4915 - 0.0396I + 0.0672I^2. \tag{8}$$

At 4.0 m, $\Delta = 2.7$, and the fitted model becomes:

$$T(I) = 1.5580 - 0.6296I + 0.1512I^2. \tag{9}$$

These depth-specific slices show that the fitted relationship is approximately linear at the focal plane, while curvature becomes stronger as depth increases. This supports the interpretation that, in the Dynamic VAC condition, increasing distance from the focal plane does not simply add a constant time penalty but also changes how strongly movement time scales with task difficulty.

To further show how the proposed model captures the behavior of the Dynamic VAC data, we compared the model predictions with the observed movement times at some depths in the Dynamic VAC data

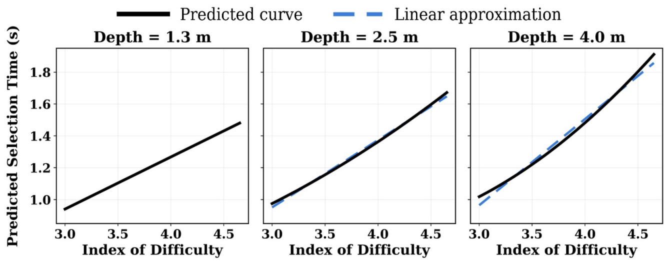


Figure 6: Model-based depth slices derived from the fitted quadratic depth-curvature model for the Dynamic VAC condition. Each panel shows predicted movement time as a function of the index of difficulty at a fixed depth. Near the focal plane, the fitted relationship is close to linear, while at farther depths it increasingly curves.

(Fig. 7). The plots show the relationship between movement time and index of difficulty for depths within ± 0.10 m of 1.5 m, 2.0 m, and 4.0 m. The model predictions closely follow the observed data points with $R^2 = 0.936$ at 1.5 m, $R^2 = 0.959$ at 2.0 m, and $R^2 = 0.955$ at 4.0 m. These examples show that the proposed formulation can accurately capture the relationship between task difficulty and movement time across different depth levels in the Dynamic VAC condition.

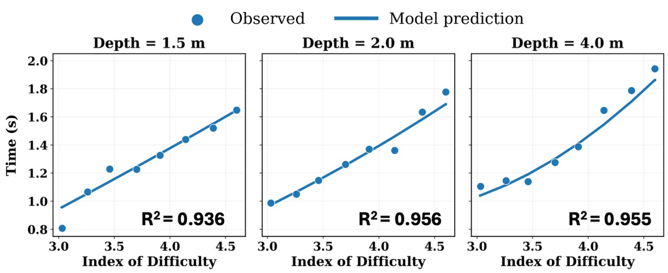


Figure 7: Example fits of the proposed depth-dependent model to the Dynamic VAC data at three depths. Points represent the observed mean movement times for each index of difficulty (ID), and the lines show the model prediction. The results show that the model captures the increasing relationship between ID and movement time across different depth levels.

As a consistency check, we evaluated how the model derived from the Dynamic VAC condition relates to the static No VAC and Constant VAC conditions. For this, we used fixed depths corresponding to the No VAC (1.3 m) and Constant VAC (4.0 m) conditions in the fitted model and compared the predicted movement times with the observed mean times across ID levels. Figure 8 shows the resulting comparisons. In both static conditions, the observed data showed a similar overall relationship between the ID and movement time as predicted by the model. Yet, the observed times in the static conditions are consistently lower than the predictions derived from the Dynamic VAC model.

Overall, this comparison suggests that the model derived for the Dynamic VAC overestimates movement time for stationary targets, but it captures the general relationship between task difficulty and performance across depth.

5.4 Subjective Measures

Cognitive Load Based on the NASA TLX scores, perceived workload differed significantly across VAC conditions (Fig. 9a). Median Overall workload scores were 27.33 for the *No VAC* condition, 28.50 for *Constant VAC*, 35.25 for *Varying VAC*, and 42.50 for *Dynamic VAC*. A Friedman test revealed a significant effect of VAC condition on perceived overall workload, $\chi^2(3) = 31.82$, $p < .001$. Post hoc Wilcoxon signed-rank tests with Bonferroni correction showed that workload was significantly higher in the *Dynamic VAC* condition than in *No VAC* ($p < .001$), *Constant VAC* ($p < .001$), and *Varying VAC* ($p < .001$). Additionally, the *Varying VAC* condition produced significantly higher

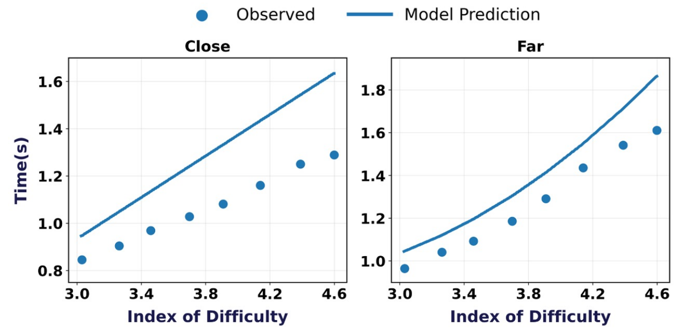


Figure 8: Comparison between observed movement times and predictions from the Dynamic VAC model for the static depth conditions (left: No VAC, target depth = 1.3 m; right: Constant VAC, target depth = 4.0 m). In both cases, the model captures the increasing relationship between the index of difficulty and movement time, while the observed static-condition times are consistently lower because targets do not change depth during acquisition.

workload than *No VAC* ($p = .039$). No significant differences were observed between *No VAC* and *Constant VAC*, or between *Constant VAC* and *Varying VAC*. The individual NASA-TLX subscale results are shown in Fig. 9. Across the six subscales, several significant pairwise differences were observed:

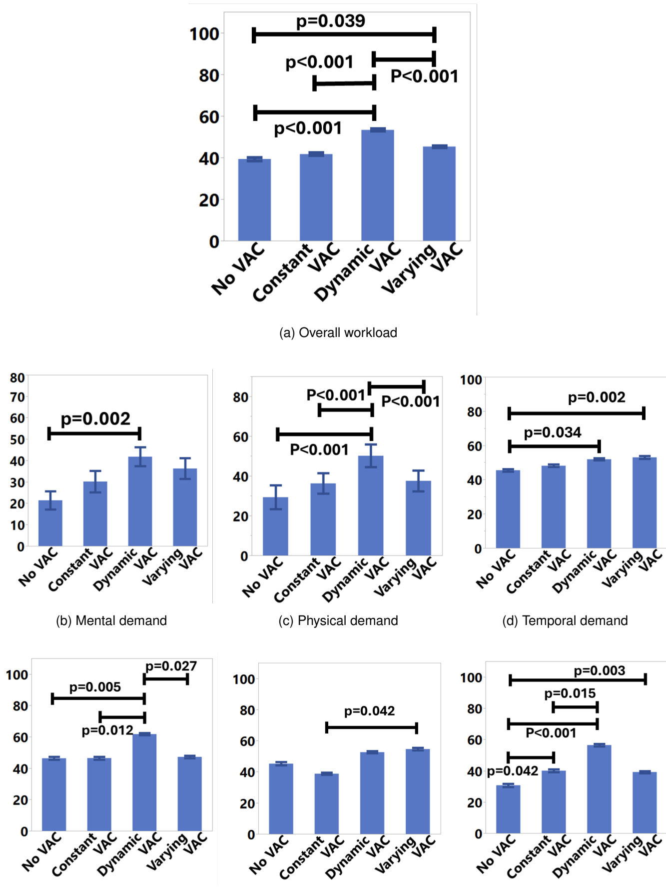
- **Mental Demand:** *Dynamic VAC* produced significantly higher mental demand than *No VAC* ($p = 0.002$).
- **Physical Demand:** *Dynamic VAC* resulted in significantly higher physical demand than *No VAC* ($p < 0.001$), *Constant VAC* ($p < 0.001$), and *Varying VAC* ($p < 0.001$).
- **Temporal Demand:** In both the *Dynamic VAC* ($p = 0.034$) and *Varying VAC* ($p = 0.002$) conditions temporal demand was significantly higher compared to the *No VAC* condition.
- **Performance:** Participants reported lower perceived performance in the *Dynamic VAC* condition compared to *No VAC* ($p = 0.005$), *Constant VAC* ($p = 0.012$), and *Varying VAC* ($p = 0.027$).
- **Effort:** Effort ratings were significantly higher in *Varying VAC* than in *Constant VAC* ($p = 0.042$).
- **Frustration:** *Dynamic VAC* produced significantly higher frustration than *No VAC* ($p < 0.001$) and *Constant VAC* ($p = 0.015$). In addition, both *Constant VAC* ($p = 0.042$) and *Varying VAC* ($p = 0.003$) produced higher frustration than *No VAC*.

System Usability Scale (SUS). A Friedman test did not reveal a significant effect of VAC condition on perceived usability, $\chi^2(3) = 4.53$, $p = .209$. As shown in Table 3, the mean SUS scores for all four conditions remained within the *Good* usability range. Median SUS scores were 80.00 for the *No VAC* condition, 77.50 for *Constant VAC*, 75.00 for *Varying VAC*, and 77.50 for *Dynamic VAC*. Overall, these results suggest that participants did not perceive system usability as significantly different across the four VAC conditions. This indicates that the experimental setup itself was consistent and that differences observed in performance and workload measures are unlikely to be driven by interface usability.

Table 3: SUS Results and Grades.

Technique	SUS Score (M)	SUS Grade
No VAC	79.8	B (Good)
Constant VAC	77.4	B (Good)
Dynamic VAC	77.1	B (Good)
Varying VAC	75.2	B (Good)

Post-study feedback. At the end of the experiment, participants answered several questions regarding fatigue, visual discomfort, and their visual perception of the targets. Responses indicated varying



(a) Overall workload

(b) Mental demand

(c) Physical demand

(d) Temporal demand

(e) Performance

(f) Effort

(g) Frustration

Figure 9: NASA-TLX results across the four VAC conditions. The top panel shows the overall NASA-TLX raw workload score. The lower panels show the six NASA-TLX subscales. Error bars indicate variability across participants. Significant pairwise comparisons are shown above the bars.

levels of physical and mental fatigue on the 7-point scales, and 9/24 participants reported experiencing eye discomfort during the study (12/24 reported none, 3/24 were not sure). Reports of perceived inconsistencies in the visual presentation were limited: 3/24 participants reported noticing inconsistencies in target depth, and 2/24 reported that targets appeared to move or shift unintentionally. Reports of motion sickness symptoms were rare (1/24 participants). Overall, these responses suggest that while the task imposed noticeable visual and attentional demands, participants generally perceived the visual presentation and target behavior as stable throughout the experiment.

6 Discussion

In this paper, we examined how the vergence–accommodation conflict (VAC) affects user performance when target depth continuously changes rather than remaining fixed. We replicated and implemented the ISO 9241-411 multidirectional selection task in VR while keeping angular target size and angular target distance constant. As a result, the targets subtended the same retinal size even though their virtual depth changed.

6.1 Dynamic Target Selection

The motor performance results show that the Dynamic VAC condition significantly degraded performance compared to the No VAC condition (i.e., when targets were placed at the headset's focal plane). Participants required more time to acquire the targets in the Dynamic VAC condition, indicating slower movement execution, and also produced a significantly higher error rate than in the No VAC condition, suggesting greater difficulty in successfully selecting the intended targets. Accuracy also decreased, reflected by larger deviations from the target center at selection. Consequently, throughput—a combined measure of speed and accuracy—was lower under the Dynamic VAC, indicating an overall reduction in interaction efficiency. Overall, the hypothesis was partially supported. The Dynamic VAC condition produced significantly higher error rates than all stationary-target conditions and significantly longer movement times than the No VAC condition.

Beyond confirming the performance cost of the VAC [16], an important result of this study is that the *Dynamic VAC* condition consistently produced the lowest overall performance among the tested conditions. Unlike the static depth conditions, where vergence demand remains stable during the movement, a Dynamic VAC introduces a continuously changing vergence demand while accommodation remains fixed at the headset focal plane. In this situation, visual depth perception is not constant during the interaction. Instead, the relationship between vergence and accommodation evolves throughout the target acquisition process. This means that participants must perform the motor task while the visual conditions underlying depth perception are continuously changing.

This difference is important because it suggests that the *Dynamic VAC* condition is not simply a stronger version of a static VAC. In the *Constant VAC* condition, the visual mismatch remains stable during the pointing movement. In the *Varying VAC* condition, depth changes occur across selections but not during the movement itself. In contrast, the *Dynamic VAC* condition introduces a temporally evolving visual constraint while the participant performs the selection movement. As a result, the user must simultaneously find the target (most importantly in the continuously changing depth dimension) and point to it, and adapt to the changing vergence demand. This additional instability may have contributed to the increased movement time and higher error rate. Past work has shown that typical speed ranges for human vergence motion-in-depth are slow, often only a few degrees per second or even in the tenths of a degree per second [45], which may explain this poorer performance.

While Dynamic VAC produced the lowest overall performance, the detailed pattern across metrics provides additional insight into how participants adapted to the task. Angular endpoint variability (SD_x) was highest in the Varying VAC condition rather than in the Dynamic condition. One explanation is that participants adapted a more cautious movement strategy in the Dynamic VAC condition, slowing down to compensate for the changing visual constraint. While this strategy may have partially lowered endpoint variability, it did not fully compensate for the increased perceptual demands, resulting in overall slower movements and reduced throughput.

6.2 Dynamic Vergence-Accommodation Conflict

Our findings are consistent with prior research showing that the VAC affects interaction performance in VR [13,15,16,55]. Previous work has reported slower movement times, reduced throughput, and decreased pointing performance when targets are placed away from the focal plane or when multiple depth planes are involved. Our results align with these observations, but extend them by examining a condition where depth changes during the movement itself. While prior work primarily focused on static or discretely varying depth configurations, our study shows that continuously changing depth during acquisition introduces additional performance costs that are not fully captured by static depth evaluations. Although our study was conducted in VR, these findings may also be relevant to optical see-through AR interactions, where users switch between real and virtual content at different focal depths [5,6].

The depth-dependent modeling results provide further insight into this behavior. In traditional Fitts' law analyses, movement time is modeled primarily as a function of task difficulty (ID). However, in the Dynamic VAC condition, depth becomes an additional factor affecting performance. The 3D movement-time surface model shows that movement time increases with both task difficulty and distance from the focal plane. Also, the relationship between ID and movement time does not remain linear as depth increases. Instead, the curvature of the

ID–time relationship becomes stronger at greater depths, suggesting that the effect of task difficulty is modulated by the combined changes in target depth and vergence demand.

This behavior is captured by the quadratic depth–curvature model fitted to the Dynamic VAC data. Near the focal plane, the predicted relationship between ID and movement time remains approximately linear, which is consistent with classical Fitts' law behavior. However, as the target moves farther from the focal plane, the model predicts increasing curvature in the ID–time relationship. In practical terms, this means that depth not only adds a time cost but also changes how strongly performance is affected by task difficulty. The resulting surface therefore provides a more complete representation of interaction performance when targets move in depth.

The comparison between the Dynamic-derived model and the static No VAC (close targets at 1.3 m) and Constant VAC (far targets at 4.0 m) conditions further supports this interpretation. When evaluated at fixed depths corresponding to the static conditions, the model preserves the overall relationship between ID and movement time observed in the Dynamic VAC data. However, the predicted movement times are consistently higher than those measured in the conditions where the target is stationary. This difference is expected because the Dynamic condition includes an additional challenge: targets continuously change depth during acquisition, and users have to adapt to that change, while stationary targets remain stable. The comparison, therefore, suggests that the Dynamic VAC model captures the underlying ID–depth relationship while incorporating the additional cost introduced by depth motion during interaction.

In this work, we collected 9_{ID} x 11 targets x 3 repetitions x 24 participants, resulting in 7128 data points for the Dynamic VAC condition. Since the target depth at the time of selection was not constrained, data were distributed continuously across the depth range rather than accumulating at specific distances. A distribution analysis also confirmed that there were at least 180 observations within each 10 cm depth interval.

The subjective measures provide complementary context for the objective results. NASA TLX scores indicated that participants perceived a higher workload in the Dynamic VAC condition, which is consistent with the slower movements and higher error rates observed in the task. At the same time, SUS scores did not differ significantly across conditions, suggesting that participants perceived the interaction technique itself as similarly usable across all conditions. This observation indicates that the differences in performance are unlikely to be caused by interface usability differences.

6.3 Design Implications

From a design perspective, our results provide several implications for interaction techniques and interfaces in VR.

• Continuously changing depth introduces additional perceptual and motor coordination demands compared to stationary targets. Designers should therefore **avoid requiring high-precision selection or confirmation while targets are moving along the depth axis**, particularly when they are far from the headset's focal plane.

• The proposed depth-dependent performance model shows that depth not only increases movement time but also alters how performance scales with task difficulty. This suggests that proposed 3D UI **interaction techniques should also be evaluated at different depths with different task difficulties**.

• Increased NASA-TLX results in the Dynamic VAC condition, suggesting that depth motion introduces additional cognitive and perceptual loads. Thus, we recommend **minimizing task complexity when the users interact with the targets moving in depth**.

• Many real-world VR scenarios involve objects that approach or move away from the user, such as training simulations [36] and rhythm and action games [17]. As the interaction performance measurement with stationary targets may underestimate the task difficulty, **evaluations should be based on the depth-varying conditions to better reflect realistic usage contexts**.

6.4 Limitations

The experiment was conducted using a single commercial VR HMD with a fixed focal plane. Since the VAC's magnitude depends on the optical properties of the display, the observed effects may differ across headsets with different focal distances or optical designs.

Although the Meta Quest Pro supports adjustable lens spacing, headset-specific optical and stereo-rendering characteristics may still have affected the results. Residual radial lens distortion can affect how users perceive surface orientation [56], while mismatches among a user's eye positions, the headset optics, and the virtual rendering cameras can distort perceived surface orientation and distance [57, 63]. Chromatic aberration and other headset-specific optical artifacts may also vary across the optical field and affect visual performance. These factors were not independently measured in our study and may therefore have contributed to the observed effects in addition to the VAC.

Second, we used a controlled ISO 9241-411 selection paradigm, which is appropriate for isolating performance effects but does not represent the full range of interaction scenarios found in VR applications. Third, in the Dynamic VAC condition, target depth and vergence demand changed together. Therefore, their individual contributions cannot be separated in our single-focal setup, and the results should be interpreted as reflecting the combined effects of depth motion and changing VAC rather than VAC alone. Although target scaling maintained a constant visual angle and retinal size across depths, it did not guarantee constant perceived size. Perceived size may still vary with vergence-defined depth, which may have contributed to the observed performance differences. Also, changing the speed of the targets' movement might change the effects of the VAC and the results here.

Future work could examine how the Dynamic VAC affects other interaction techniques, such as virtual-hand interaction or gaze-based selection, as well as tasks involving continuous confirmation rather than discrete button presses. Additional studies could investigate nonlinear and variable-speed depth trajectories, as well as lateral target motion at fixed depths, to better separate the effects of depth motion and changing VAC. Further, the proposed model should be evaluated on unseen motion conditions and at larger viewing distances to determine whether the effects observed in the Dynamic VAC condition remain measurable beyond the distances typically examined in VAC studies.

Overall, our results show that continuously changing target depth introduces an additional challenge for 3D interaction in single-focal VR head-mounted displays. Understanding how interaction performance changes when depth evolves during the movement is important for accurately evaluating VR interfaces and for designing interactive systems that involve targets moving in depth.

7 Conclusion

This paper examined how continuously changing target depth affects 3D selection performance in Virtual Reality with single-focal head-mounted displays. Unlike prior work that focused on static depth conditions, we introduced a *Dynamic VAC* condition in which targets changed depth during selection and compared it with No VAC, Constant VAC, and Varying VAC conditions in an ISO 9241-411 task. Results from a within-subjects study with 24 participants showed that a Dynamic VAC reduced interaction efficiency, leading to slower movements, higher error rates, and lower throughput than the No VAC condition. Our findings also showed that movement time was affected not only by task difficulty but also by distance from the focal plane, with the relationship between the index of difficulty and movement time becoming increasingly non-linear at greater depths. Overall, these results suggest that targets moving in depth create distinct visual and motor coordination challenges, highlighting the importance of accounting for dynamic depth changes when evaluating performance and designing VR interfaces.

References

[1] Ergonomics of human–system interaction — part 411: Evaluation methods for the design of physical input devices, 2015. 3, 4

[2] H. Adams, J. Stefanucci, S. Creem-Regehr, and B. Bodenheimer. Depth perception in augmented reality: The effects of display, shadow, and

position. In *2022 IEEE conference on virtual reality and 3D user interfaces (VR)*, pp. 792–801. IEEE, 2022. doi: 10.1109/VR51125.2022.00101 2

[3] H. Akaike. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6):716–723, 1974. doi: 10.1109/TAC.1974.1100705 5, 6

[4] V. Angelov, E. Petkov, G. Shipkovenski, and T. Kalushkov. Modern virtual reality headsets. In *2020 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)*, pp. 1–5, 2020. doi: 10.1109/HORA49412.2020.9152604 1

[5] M. S. Arefin, N. Phillips, A. Plopski, J. L. Gabbard, and J. E. Swan. The effect of context switching, focal switching distance, binocular and monocular viewing, and transient focal blur on human performance in optical see-through augmented reality. *IEEE Transactions on Visualization and Computer Graphics*, 28(5):2014–2025, 2022. doi: 10.1109/TVCG.2022.3150503 2, 8

[6] D. Au, R. Allison, and L. Wilcox. Depth cue conflicts between real and virtual objects disrupt visually-guided reaching in augmented reality. *IEEE Transactions on Visualization and Computer Graphics*, 32:3552–3561, 04 2026. doi: 10.1109/TVCG.2026.3679915 2, 8

[7] D. Au, R. S. Allison, I. Gunasekera, and L. M. Wilcox. Misperception of the distance of virtual augmentations. In *2025 IEEE Conference Virtual Reality and 3D User Interfaces (VR)*, pp. 516–525, 2025. doi: 10.1109/VR59515.2025.00075 1

[8] M. S. Banks, J. Kim, and T. Shibata. Insight into vergence-accommodation mismatch. In *Head- and Helmet-Mounted Displays XVIII: Design and Applications*, vol. 8735 of *Proceedings of SPIE*, p. 873509, 2013. doi: 10.1117/12.2019866 2, 3

[9] M. D. Barrera Machuca and W. Stuerzlinger. Do stereo display deficiencies affect 3d pointing? In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI EA '18, pp. 1–6. Association for Computing Machinery, New York, NY, USA, 2018. doi: 10.1145/3170427.3188540 2

[10] 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, pp. 1–14. Association for Computing Machinery, New York, NY, USA, 2019. doi: 10.1145/3290605.3300437 2, 3

[11] M. R. Bashar, M. D. Barrera Machuca, W. Stuerzlinger, and A. U. Batmaz. The effect of visual depth on the vergence–accommodation conflict on 3d selection performance within virtual reality headsets. *The Visual Computer*, pp. 1–17, 2025. doi: 10.1007/s00371-025-03990-x 2, 3, 4, 5, 6

[12] 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 *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, CHI '22, art. no. 633, 15 pp. Association for Computing Machinery, New York, NY, USA, 2022. doi: 10.1145/3491102.3502067 2, 3

[13] A. U. Batmaz, M. Hudhud Mughrabi, M. D. Barrera Machuca, and W. Stuerzlinger. Effect of stereo deficiencies on virtual distal pointing. In *Proceedings of the 28th ACM Symposium on Virtual Reality Software and Technology*, VRST '22, art. no. 12, 8 pp. Association for Computing Machinery, New York, NY, USA, 2022. doi: 10.1145/3562939.3565621 2, 3, 4, 8

[14] A. U. Batmaz, M. Hudhud Mughrabi, M. Sarac, M. D. Barrera Machuca, and W. Stuerzlinger. Measuring the effect of stereo deficiencies on peripersonal space pointing. In *Conference on Virtual Reality and 3D User Interfaces*, VR '23, 11 pp., Mar 2023. doi: 10.1109/VR55154.2023.00063 2, 3

[15] 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 2, 8

[16] A. U. Batmaz, R. Turkmen, M. Sarac, M. D. Barrera Machuca, and W. Stuerzlinger. Re-investigating the effect of the vergence-accommodation conflict on 3d pointing. In *Proceedings of the 29th ACM Symposium on Virtual Reality Software and Technology*, VRST '23, art. no. 8, 10 pp. Association for Computing Machinery, New York, NY, USA, 2023. doi: 10.1145/3611659.3615686 2, 3, 4, 8

[17] Beat Games. Beat saber, 2019. Virtual reality rhythm game. 9

[18] D. Bowman, C. Wingrave, J. Campbell, and V. Ly. Using pinch gloves for both natural and abstract interaction techniques in virtual environments. *Proceedings of HCI International*, 01 2001. 4

[19] J. Brooke. SUS: A 'Quick and Dirty' Usability Scale. In *Usability Evaluation In Industry*. CRC Press, London, 1996. 4

[20] J. Cashion, C. Wingrave, and J. J. LaViola Jr. Dense and dynamic 3d selection for game-based virtual environments. *IEEE Transactions on Visualization and Computer Graphics*, 18(4):634–642, 2012. doi: 10.1109/TVCG.2012.40 3

[21] 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, 2013. doi: 10.1016/j.ergon.2013.05.005 3

[22] 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 3

[23] J. I. Cross, C. Boag-Hodgson, and T. J. Mavin. Measuring presence and situational awareness in a virtual reality flight simulator. *Aviation Psychology and Applied Human Factors*, 13(2):83–94, 2023. 1

[24] J. E. Cutting and P. M. Vishton. Perceiving Layout and Knowing Distances: The Integration, Relative Potency and Contextual Use of Different Information about Depth. In *Handbook of Perception and Cognition, Vol 5; Perception of Space and Motion*, pp. 69–117. Academic Press, San Diego, California, USA, 1995. doi: 10.1016/B978-012240530-3/50005-5 1

[25] F. Daniel and Z. Kapoula. Induced vergence-accommodation conflict reduces cognitive performance in the stroop test. *Sci. Rep.*, 9(1):1247, Feb. 2019. 2

[26] G. de Haan, M. Koutek, and F. H. Post. Intenselect: using dynamic object rating for assisting 3d object selection. In *Proceedings of the 11th Eurographics Conference on Virtual Environments*, EGVE'05, pp. 201–209. Eurographics Association, Goslar, DEU, 2005. 3

[27] S. Durrani, C. Onyedimma, R. Jarrah, A. Bhatti, K. R. Nathani, A. R. Bhandarkar et al. The virtual vision of neurosurgery: How augmented reality and virtual reality are transforming the neurosurgical operating room. *World Neurosurgery*, 168:190–201, 2022. doi: 10.1016/j.wneu.2022.10.002 1

[28] P. M. Fitts. The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, 47(6):381–391, 1954. 3, 5, 6

[29] 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 2

[30] J. F. Hair Jr, W. C. Black, B. J. Babin, and R. E. Anderson. Multivariate data analysis, 2014. 4

[31] S. G. Hart. Nasa-task load index (nasa-tlx); 20 years later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 50(9):904–908, 2006. doi: 10.1177/154193120605000909 4

[32] 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):33–33, 03 2008. doi: 10.1167/8.3.33 2, 3

[33] X. Hu, M. Perusquía-Hernández, M. D. B. Machuca, A. U. Batmaz, Y. Zhang, W. Stuerzlinger et al. Varifocal displays reduce the impact of the vergence-accommodation conflict on 3d pointing performance in augmented reality systems. https://arxiv.org/abs/2602.05129, 2026. 2

[34] J. Huang, J. J. Dudley, S. Uzor, D. Wu, P. O. Kristensson, and F. Tian. Understanding user performance of acquiring targets with motion-in-depth in virtual reality. *International Journal of Human-Computer Studies*, 163:102817, 2022. doi: 10.1016/j.ijhcs.2022.102817 3

[35] 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. Association for Computing Machinery, New York, NY, USA, 2016. doi: 10.1145/2858036.2858244 2

[36] E. Karatas, K. Sunday, S. E. Apak, Y. Li, J. Sun, A. U. Batmaz et al. "ï consider vr table tennis to be my secret weapon!": An analysis of the vr table tennis players' experiences outside the lab". In *Proceedings of the 2023 ACM Symposium on Spatial User Interaction*, SUI '23, art. no. 29, 12 pp. Association for Computing Machinery, New York, NY, USA, 2023. doi: 10.1145/3607822.3614539 1, 9

[37] D. Kenoyer-Healy, M. Sturgeon, and M. S. Arefin. Occlusion effect on action space distance estimations in optical see-through and video see-through augmented reality. *IEEE Transactions on Visualization and Computer Graphics*, 32(5):3977–3987, 2026. doi: 10.1109/TVCG.2026.3680623 2

[38] F. L. Kooi and A. Toet. Visual comfort of binocular and 3d displays. *Displays*, 25(2):99–108, 2004. doi: 10.1016/j.displa.2004.07.004 2

[39] 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. doi: 10.1016/j.ijhcs.2010.05.001 2, 3

[40] P. B. Kruger, K. R. Aggarwala, S. Bean, and S. Mathews. Accommodation to stationary and moving targets. *Optometry and Vision Science*, 74(7):505–510, 1997. doi: 10.1097/00006324-199707000-00018 2

[41] I. S. MacKenzie. Fitts' law as a research and design tool in human-computer interaction. *Human–Computer Interaction*, 7(1):91–139, 1992. doi: 10.1207/s15327051hci0701_3 3, 5, 6

[42] K. Mackenzie, D. Hoffman, and S. Watt. Accommodation to multiple-focal-plane displays: Implications for improving stereoscopic displays and for accommodation control. *Journal of Vision*, 10:22–22, 07 2010. doi: 10.1167/10.8.22 2

[43] P. Mallery and D. George. SPSS for windows step by step: a simple guide and reference. *Allyn, Bacon, Boston,*, 2003. 4

[44] A. Murata and H. Iwase. Extending fitts' law to a three-dimensional pointing task. *Human Movement Science*, 20(6):791–805, 2001. doi: 10.1016/S0167-9457(01)00058-6 2, 3

[45] H. T. Nefs and J. M. Harris. Vergence effects on the perception of motion-in-depth. *Experimental Brain Research*, 183(3):313–322, 2007. doi: 10.1007/s00221-007-1046-5 2, 3, 8

[46] H. T. Nefs, L. O'Hare, and J. M. Harris. Two independent mechanisms for motion-in-depth perception: Evidence from individual differences. *Frontiers in Psychology*, Volume 1 - 2010, 2010. doi: 10.3389/fpsyg.2010.00155 2

[47] K. Pfeil, S. Masnadi, J. Belga, J.-V. T. Sera-Josef, and J. LaViola. Distance perception with a video see-through head-mounted display. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pp. 1–9, 2021. doi: 10.1145/3411764.3445223 2

[48] D. Regan, S. J. Hamstra, S. Kaushal, A. Vincent, R. Gray, and K. I. Beverley. Visual processing of the motion of an object in three dimensions for a stationary or a moving observer. *Perception*, 24(1):87–103, 1995. doi: 10.1068/p240087 2, 3

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

[50] G. E. Schwarz. Estimating the dimension of a model. *The Annals of Statistics*, 6(2):461–464, 1978. 5, 6

[51] N. E. Seymour, A. G. Gallagher, S. A. Roman, M. K. O'brien, V. K. Bansal, D. K. Andersen et al. Virtual reality training improves operating room performance: results of a randomized, double-blinded study. *Annals of surgery*, 236(4):458–464, 2002. 1

[52] R. W. Soukoreff and I. S. MacKenzie. Towards a standard for pointing device evaluation, perspectives on 27 years of fitts' law research in hci. *International Journal of Human-Computer Studies*, 61(6):751–789, 2004. Fitts' law 50 years later: applications and contributions from human-computer interaction. doi: 10.1016/j.ijhcs.2004.09.001 3

[53] D. P. Spiegel and I. M. Erkelens. Stereopsis–occlusion conflicts impair visual performance in augmented reality. *Journal of the Society for Information Display*, 33(8):937–947, 2025. doi: 10.1002/jsid.2095 1

[54] A.-N. Spiess and N. Neumeyer. An evaluation of r2 as an inadequate measure for nonlinear models in pharmacological and biochemical research: a monte carlo approach. *BMC pharmacology*, 10(1):6, 2010. 5, 6

[55] R. J. Teather and W. Stuerzlinger. Pointing at 3d targets in a stereo head-tracked virtual environment. In *Proceedings of the 2011 IEEE Symposium on 3D User Interfaces*, 3DUI '11, pp. 87–94. IEEE Computer Society, USA, 2011. 2, 3, 8

[56] J. Tong, R. Allison, and L. Wilcox. The impact of radial distortions in vr headsets on perceived surface slant. *Journal of Imaging Science and Technology*, 63, 11 2019. doi: 10.2352/J.ImagingSci.Technol.2019.63.6.060409 9

[57] J. Tong, L. M. Wilcox, and R. S. Allison. The impacts of lens and stereo camera separation on perceived slant in virtual reality head-mounted displays. *IEEE Transactions on Visualization and Computer Graphics*, 28(11):3759–3766, 2022. doi: 10.1109/TVCG.2022.3203098 9

[58] C. Vienne, S. Masfrand, C. Bourdin, and J.-L. Vercher. Depth perception in virtual reality systems: effect of screen distance, environment richness and display factors. *IEEE Access*, 8:29099–29110, 2020. doi: 10.1109/ACCESS.2020.2972122 2

[59] 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 2

[60] X. M. Wang, D. Southwick, I. Robinson, M. Nitsche, G. Resch, A. Mazalek et al. The geometry of the vergence-accommodation conflict in mixed reality systems. *Virtual Reality*, 28(2):95, 2024. 2

[61] 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 4

[62] Y. Zheng, J. Huang, H. Zhang, Y. Bian, J. Liu, C. Yang et al. 3d ternary-gaussian model: Modeling pointing uncertainty of 3d moving target selection in virtual reality. *International Journal of Human-Computer Studies*, 198:103454, 2025. doi: 10.1016/j.ijhcs.2025.103454 3

[63] R. X. Zhu, C. S. Burlingham, O. Mercier, and P. Guan. Perceptual sensitivity to stereo geometry errors in head-mounted displays, 2026. 9