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

Zahra Rasoulifar

A Thesis

in

The Department

of

Computer Science and Software Engineering

Presented in Partial Fulfillment of the Requirements

for the Degree of

Master of Applied Computer Science (Computer Science) at

Concordia University

Montréal, Québec, Canada

April 2026

CONCORDIA UNIVERSITY

School of Graduate Studies

This is to certify that the thesis prepared

By: **Zahra Rasoulifar**

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

and submitted in partial fulfillment of the requirements for the degree of

Master of Applied Computer Science (Computer Science)

complies with the regulations of this University and meets the accepted standards with respect to originality and quality.

Signed by the Final Examining Committee:

______________________________________ Chair
Dr. Charalambos Poullis

______________________________________ Examiner
Dr. Mahdi Hosseini

______________________________________ Supervisor
Dr. Anil Ufuk Batmaz

Approved by ____
Joey Paquet, Chair
Department of Computer Science and Software Engineering

_______________ 2026

____
Mourad Debbabi, Dean
Faculty of Engineering and Computer Science

Abstract

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

Zahra Rasoulifar

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.

Acknowledgments

I would like to sincerely thank my supervisor, Dr. Anil Ufuk Batmaz, for his continuous guidance, support, and encouragement throughout this work. His feedback and perspective shaped this research in meaningful ways, and I am truly grateful for his mentorship.

I am also thankful to my co-authors, Dr. Wolfgang Stuerzlinger and Dr. Mayra D. Barrera Machuca, for their collaboration and contributions to this work.

A special thanks to Rumeysa Turkmen for her constant support, guidance, and for always being there to help, both academically and beyond. I would also like to thank the other members of the EXIT Lab for their support, discussions, and help throughout this process.

Finally, I am deeply grateful to my family and friends for their support, encouragement, and patience throughout this process, and for the constant presence of those who made this easier and more meaningful.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Extended Reality (XR) is an umbrella term that refers to a range of immersive computing technologies that combine real and virtual environments within a unified interaction space [7]. XR includes Virtual Reality (VR), Augmented Reality (AR), and Mixed Reality (MR), which differ in terms of how digital content is integrated with the physical world, as illustrated along the Milgram's [1] reality–virtuality continuum in Figure 1.1. These technologies aim to move beyond traditional screen-based interaction by enabling users to perceive, manipulate, and navigate digital information through input modalities [8], such as controllers, hand gestures, gaze, speech and voice commands [9], and body movement. XR systems are typically experienced through dedicated hardware platforms, including head-mounted displays (HMDs), CAVE systems, mobile devices, or spatial computing glasses, and are increasingly used in applications such as training simulation [10], engineering design [11], healthcare [12], education [13], and collaborative visualization [14].

A Head-Mounted Display (HMD) is a wearable visual interface device that is positioned directly in front of the user's eyes to deliver stereoscopic or monoscopic digital imagery [15]. HMDs are a fundamental hardware component in many XR systems because they provide controlled visual presentation and support immersive perception. From a technical perspective, an HMD integrates near-eye displays, optical lenses for focal adjustment and field-of-view expansion, inertial measurement units (IMUs) such as accelerometers and gyroscopes for head-motion tracking, and often outward-facing cameras for inside-out positional tracking and environment sensing (Figure 1.2) [7, 15]. This sensing pipeline enables real-time viewpoint updates based on user movement, which

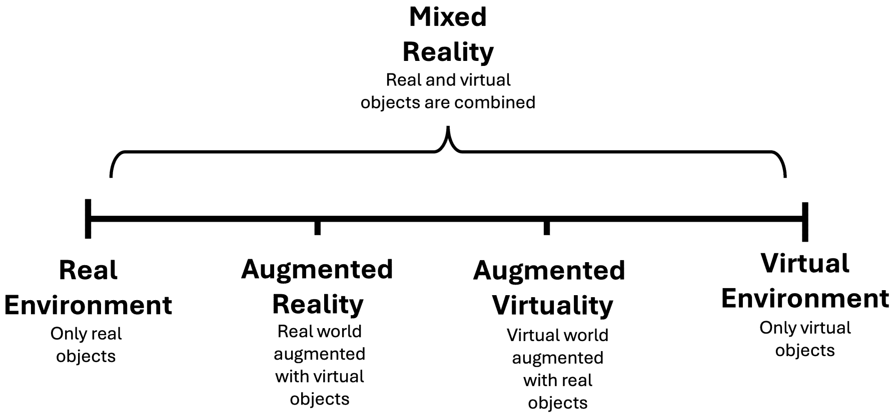


Figure 1.1: Milgram's Reality–Virtuality Continuum [1], illustrating the spectrum from real environments to fully virtual environments, including augmented and mixed reality.

is essential for maintaining spatial consistency and presence. Modern HMDs may also incorporate advanced sensing and interaction capabilities, including hand tracking, eye tracking, depth sensing, and controller input, allowing multimodal interaction within virtual or augmented environments. An example HMD is shown in Figure 1.3, which is also used in the user study included in this thesis.

Virtual Reality (VR) is a type of XR technology that generates a fully computer-simulated environment which replaces the user's perception of the physical world. In VR systems, the HMD occludes real-world visual input and presents a stereoscopic virtual scene that responds dynamically to head and body motion. Users interact with virtual objects through input devices such as handheld controllers [7], tracked hands [8], or gaze-based interfaces [8, 16]. From an engineering standpoint, VR systems require low-latency tracking, high frame rates, and accurate spatial rendering to maintain immersion and reduce discomfort [17]. VR is widely applied in domains where controlled, repeatable, and highly immersive environments are beneficial, such as simulation-based training [18–20], virtual prototyping [21], rehabilitation [22], and immersive gaming [23].

Augmented Reality (AR) is defined as a technology that superimposes digital content onto the user's view of the real environment while preserving direct perception of physical surroundings [24]. AR systems register virtual elements such as text, graphics, or 3D models within the user's visual field, typically through hand-held displays, optical see-through HMDs, or video see-through devices [24, 25]. Technically, AR requires accurate tracking and registration algorithms to align digital content with real-world coordinates, as well as environment understanding [25] to maintain spatial

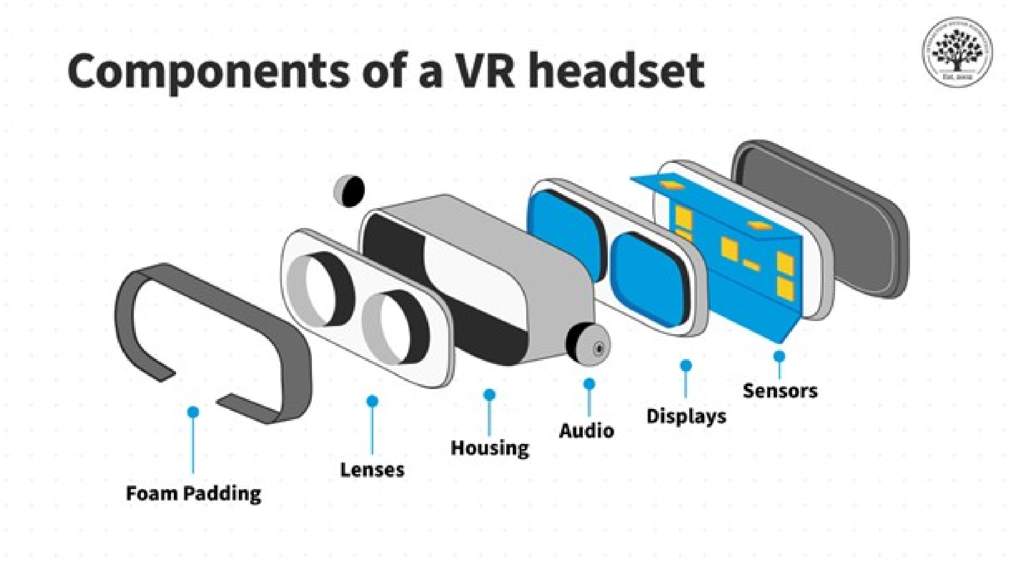


Figure 1.2: Components of a VR headset, including lenses, displays, sensors, and housing. Adapted from Interaction Design Foundation [2].

Figure 1.3: Example of a modern XR head-mounted display (Meta Quest Pro), illustrating a commercially available device used for immersive interaction. Image source: Meta Platforms, Inc. [3].

stability. AR is commonly used for applications such as navigation assistance [24, 25], industrial maintenance guidance [24], retail visualization [26], and context-aware information delivery [27].

Mixed Reality (MR) represents a more tightly integrated form of real–virtual fusion in which digital objects are not only overlaid onto the physical environment but are also spatially anchored and capable of interacting with real-world geometry. MR systems employ advanced spatial mapping, depth sensing, and real-time scene reconstruction [25] to enable virtual content to respond to surfaces, occlusions, and user movement in a physically meaningful way. This allows users to walk around holographic objects [7], manipulate them with natural gestures [8], and collaborate with others in shared hybrid environments [14]. Due to these capabilities, MR is particularly relevant for engineering visualization [11], medical planning [28], remote collaboration [14], and complex training scenarios [10] where contextual alignment between real and virtual elements is critical.

As XR systems continue to evolve, they are increasingly used in scenarios where interaction accuracy directly affects real-world outcomes [29]. In training and simulation environments, small errors in user interaction can translate to significant performance differences when skills are transferred to real-world tasks [18]. Similarly, in medical or industrial applications, imprecise interaction may negatively affect decision-making and task performance [18, 28]. Therefore, understanding the factors that affect interaction performance in XR is critical for designing reliable and effective systems [30].

Beyond application-specific requirements, XR interaction also introduces new challenges compared to traditional 2D interfaces [30]. Users must perceive and interact with virtual objects in three-dimensional (3D) space, often without the physical feedback available in real-world interactions. This makes accurate perception of spatial relationships a fundamental requirement for successful interaction. As a result, limitations in visual perception can directly affect how users perform tasks in XR environments.

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) [31, 32]. Modern XR HMDs still struggle to simulate monocular and binocular cues [33, 34]. Particularly, current commercial HMDs suffer from the vergence-accommodation conflict (VAC), 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) [34, 35]. As illustrated in Figure 1.4, vergence cues indicate a virtual depth while accommodation remains fixed at the display plane, creating a discrepancy in visual signals. This mismatch affects user performance in 3D selection tasks [36, 37]. Previous work has shown that the VAC produces a combination of increased cognitive and perceptual strain that impairs the user's ability to quickly and accurately perceive an object's spatial position [38–40], thereby affecting the speed and precision of interaction with the object [41–43].

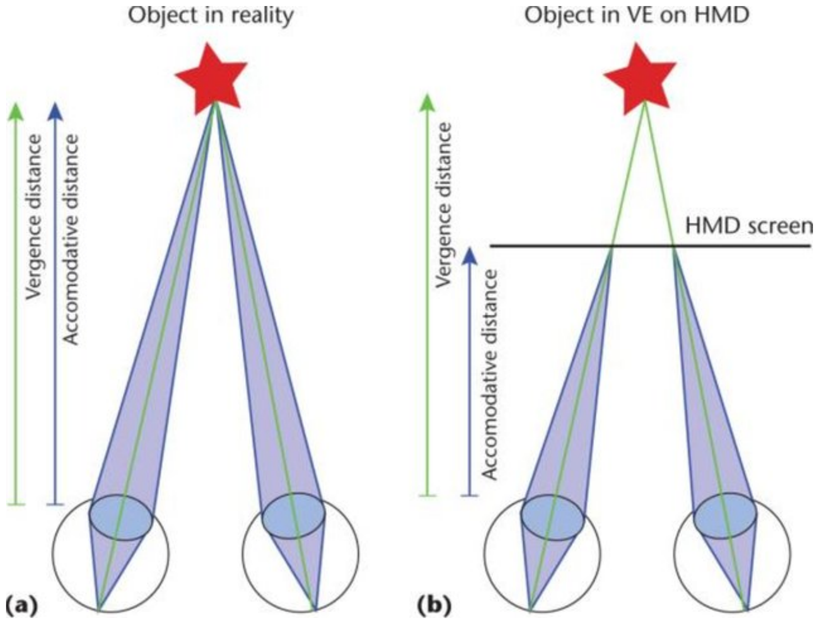


Figure 1.4: Illustration of the vergence–accommodation conflict (VAC) in XR head-mounted displays. Vergence cues suggest a virtual depth while accommodation remains fixed at the display plane.

In natural viewing conditions, vergence and accommodation are tightly coupled, meaning that both eye convergence and lens focusing adjust consistently based on the distance of the observed object [38]. However, in most XR HMDs, visual content is displayed at a fixed focal distance while binocular disparity cues simulate objects at varying depths [44]. This decoupling introduces a persistent inconsistency in visual signals, forcing the visual system to reconcile conflicting information during interaction.

This mismatch is not only a perceptual issue but also has measurable consequences on user performance. When users attempt to interact with virtual objects, especially in tasks that require precise targeting, they rely heavily on accurate depth perception [44]. Any inconsistency in depth cues can lead to errors in estimating object position, resulting in slower movements, reduced accuracy, and increased effort [38, 44]. Understanding how these perceptual limitations affect interaction is therefore essential for designing better XR systems.

Previous work has demonstrated the detrimental effects of the VAC on user interaction performance in VR and AR environments [44, 45]. However, these studies investigated tasks involving stationary targets, where object positions remained fixed at specific depths. Such experimental setups are useful for modeling user motor performance in controlled scenarios that resemble everyday interactions with objects located at constant distances. In many real-world situations, however, users interact with targets that move dynamically in depth. Examples include catching a thrown ball, a goalkeeper intercepting a soccer ball, catching a flying frisbee, reacting to a dropped object such as a slipping phone, catching keys or tools tossed by another person, or a firefighter intercepting equipment thrown from below. In these scenarios, the target position continuously changes along the depth axis, requiring users to update both motor plans and visual focus in real time. Modeling such interactions in XR environments introduces additional challenges. As the target moves in depth, the visual system must continuously adjust vergence to maintain fixation, while the accommodation demand imposed by current stereoscopic displays remains fixed. This dynamic mismatch may further affect visuomotor coordination and movement timing compared to the VAC in stationary targets. Despite the ecological relevance of these tasks, the impact of VAC under dynamic depth conditions has not been systematically investigated in prior XR interaction studies.

In this thesis, we investigate the effect of a dynamic vergence–accommodation conflict on 3D selection performance in VR. Specifically, we examine how a continuously changing VAC condition, produced when targets move toward and away from the user, affects target selection compared to fixed-depth conditions. Chapter 3 presents our study and its findings in detail.

Chapter 2

Background

2.1 Human Depth Perception and Visual Cues

Depth perception refers to the ability of the human visual system to interpret the three-dimensional structure of the environment and estimate the relative or absolute distance of objects. Rather than relying on a single source of information, the visual system integrates multiple depth cues, whose contribution depends on factors such as viewing distance, task requirements, and cue reliability [31, 32, 46]. This integration is particularly critical in near-field (peri-personal) space, where precise depth estimation directly supports visually guided motor actions such as reaching and pointing [31, 47].

Depth cues are commonly categorized into monocular and binocular cues [31, 48]. Monocular cues include relative size, occlusion, texture gradients, shading, motion parallax, and accommodation, and can be perceived with one eye [31, 48]. Binocular cues arise from the coordinated use of both eyes and include stereopsis and vergence [31, 48]. In natural viewing conditions, these cues are not processed independently; instead, they are combined to produce a coherent and stable perception of spatial layout [31]. For objects within arm's reach, accurate depth perception relies on the effective integration of multiple complementary cues, as no single cue is sufficient on its own [31].

2.1.1 Binocular Depth Cues: Vergence and Disparity

Stereopsis is a key binocular cue that arises from the horizontal disparity between the retinal images of the left and right eyes [48]. Because each eye observes the scene from a slightly different viewpoint, these disparities provide information about relative depth between objects [31]. Stereopsis is especially important for fine-grained depth judgments and supports accurate spatial discrimination in near-field interaction tasks [48].

Vergence refers to the inward or outward rotation of the eyes required to align both foveae (the regions of highest visual acuity on the retina) on a target. As an object moves closer, the eyes converge; as it moves farther away, they diverge [48]. This oculomotor response provides an additional signal related to viewing distance and plays a crucial role in depth perception within peri-personal space [32]. Vergence is also involved when observing objects moving in depth, where eye movements must continuously adjust to maintain fixation [49]. However, prior work shows that vergence can affect and bias the perception of motion-in-depth, for example by affecting the perceived speed of targets moving in depth [49]. This indicates that dynamic depth perception arises from an interaction between visual signals and oculomotor responses, rather than from a single cue alone. Such dynamic depth perception is particularly relevant in interactive environments where targets may change position along the depth axis.

2.1.2 Accommodation and Focus Mechanisms

Accommodation is the process by which the crystalline lens changes shape to maintain a sharp retinal image as viewing distance varies [50]. When focusing on nearby objects, the lens increases its optical power; when focusing on distant objects, the optical demand decreases. In natural viewing, accommodation provides a depth-related signal, particularly at short distances where changes in focal demand are more pronounced [32, 51].

Although accommodation alone is generally not sufficient for precise depth estimation, it contributes to the overall perception of depth and supports stable vision [31, 32]. Previous work has shown that accommodative responses differ between stationary and moving targets, reflecting differences in the visual stimuli driving focus adjustments [51]. This suggests that accommodative behavior depends on the properties of the visual input and operates in coordination with other depth cues during natural viewing.

2.1.3 Coupling of Vergence and Accommodation

In natural viewing conditions, vergence and accommodation are tightly coupled. When fixating on a real object, the eyes converge to its spatial location while the lens accommodates to the same physical distance. This coordinated response ensures that both disparity-based and focus-based cues provide consistent information about object depth [38, 52, 53]. As a result, the visual system can maintain single, clear vision and achieve accurate depth perception.

This coupling is not static but reflects a dynamic interaction between the oculomotor subsystems. Changes in vergence can induce accommodation responses and vice versa, demonstrating the interdependence of these mechanisms [54]. Under normal conditions, this coordinated binocular response supports efficient visually guided reaching and grasping in peri-personal space [31, 47].

Previous work has shown that, for targets in near-field viewing, depth perception relies on multiple non-pictorial cues such as stereopsis, motion parallax, vergence, and accommodation [31]. For real-world targets, these cues are consistent and jointly contribute to highly accurate depth perception [31, 32]. However, when this consistency is disrupted, depth perception can become less accurate, which has direct implications for interaction performance [38].

2.2 Vergence–Accommodation Conflict in XR Systems

2.2.1 VAC in Commercial Head-Mounted Displays

Vergence–accommodation conflict (VAC) occurs when the vergence demand specified by binocular disparity does not match the accommodative demand specified by the focal distance of the display. This situation commonly arises in stereoscopic displays used in virtual and augmented reality, where depth is simulated visually but the physical focal distance remains fixed [38, 52].

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 [38]. Related work has discussed visual comfort in binocular and 3D displays [55]. The VAC has also been connected to changes in visual system behavior (e.g., vergence and accommodation responses) when viewing stereoscopic imagery [53, 54] and to measurable costs in demanding cognitive tasks under induced conflict [40].

Current VR and AR head-mounted displays commonly operate as single-focal stereo systems, meaning that the display optics provide a fixed focal distance while binocular disparity is used to simulate depth in the virtual scene [36, 56]. As a result, accommodation is driven toward the display's focal plane, while vergence follows the depth specified by the stereo content. This mismatch between accommodative distance and vergence distance is known as the VAC [52, 56].

2.2.2 Effects of VAC on Perception and Cognition

The vergence–accommodation conflict affects perception first at the level of depth interpretation and visual comfort. When vergence demand is driven by the simulated depth of stereo content while accommodation remains tied to the display focal plane, the visual system receives inconsistent distance information. Prior work has shown that this inconsistency can reduce the accuracy of perceived depth and decrease visual performance compared to conditions in which focus cues are correct or nearly correct [38, 52]. More broadly, research on binocular and stereoscopic displays has shown that mismatches in binocular presentation can reduce visual comfort and increase viewing difficulty, especially when the visual demands are sustained or precise depth judgments are required [55]. In XR systems, these perceptual limitations are particularly important because users often need to estimate object depth accurately enough to guide interaction.

The VAC has also been linked to changes in the behavior of the oculomotor system. Vienne et al. [53] showed that vergence responses become slower under accommodation–vergence conflict than under matched focus and disparity conditions, indicating that the mismatch affects how efficiently the eyes respond to depth changes. Related work further suggests that individual differences in accommodative behavior while viewing stereoscopic stimuli are associated with the coupling between vergence and accommodation mechanisms [54]. These findings do not imply that the visual system fails completely under conflict, but they do show that the normal coordination between eye alignment and optical focus is altered. As a result, users may require more effort to achieve stable and clear viewing when interacting with stereoscopic content at depths away from the focal plane.

Table 2.1: Overview of previous 3D pointing experiments focusing on the VAC.

Publication	**Movement Direction**	**Display Device**	**Main Findings**	**Open Issues**
Lubos et al. [57]	7 directions 9 positions	Oculus Rift	Visual Perception causes errors in 3D selection	Is the VAC the cause of these errors?
Barrera and Stuerzlinger [58]	2 directions	3D TV	Depth movements are slower than lateral ones	Does this occur in the real world, too?
Barrera and Stuerzlinger [37]	2 directions	3D TV Physical apparatus	Stereo displays affect 3D pointing negatively	Generalization to other stereo displays
Batmaz et al. [36]	2 directions	HTC Vive Meta 2 (AR)	Stereo in HMDs affects 3D pointing negatively	Cause of performance drop?
Batmaz et al. [44]	2 directions	Singlefocal VR/AR bench Multifocal VR/AR bench	VAC slightly affects 3D pointing with Virtual Hand	Effect of limited reachability of targets at 70 cm
Batmaz et al. [45]	11 directions	HTC Vive 2	VAC clearly affects 3D pointing with Raycasting	How about virtual hand?
Batmaz et al. [43]	11 directions	HTC Vive 2	Constant VAC does not substantially affect 3D pointing	Instructions to participants did not emphasize speed/accuracy, effect of inconsistent grip on device
Batmaz et al. [56]	11 directions	HTC Vive 2	Biomechanics reduce the VAC's effect for virtual hand pointing	What specific biomechanical factor is responsible for this effect?
Bashar et al [41]	11 directions 6 depth distances	Oculus Quest 3	Fitts' law extension for static targets at 6 different depth distance	What is the relationship between visual depth and user selection performance?
Rasoulifar et al. (current work)	**11 directions** **Continuous depth change**	**Oculus Quest Pro**	**Fitts' law extension for target positions dynamically change in depth**	**How to model human movement when targets are not stationary?**

Beyond perception and oculomotor behavior, the VAC can also affect higher-level task performance. Daniel and Kapoula [40] showed that induced vergence–accommodation conflict increased interference effects in a Stroop task, particularly under conditions with higher cognitive demand. This suggests that the cost of resolving or tolerating the conflict is not limited to low-level vision alone, but can also affect attentional processing. In other words, the VAC can introduce an additional cognitive burden, especially in tasks that require sustained attention, visual precision, or rapid responses. Taken together, these results suggest that the VAC should not be viewed only as a display artifact or a source of discomfort. It can affect how accurately users perceive depth, how efficiently the visual system responds to stereo content, and how well users perform demanding tasks. This provides an important basis for understanding why the VAC also affects 3D interaction performance, which is discussed in the following sections.

2.2.3 Approaches to Mitigate VAC

Because the vergence–accommodation conflict originates from a mismatch between simulated depth and optical focus, many proposed solutions aim to restore a more natural relationship between these cues. In general, these approaches attempt to reduce the separation between vergence demand and accommodative demand, so that the visual system receives depth information that is more consistent with natural viewing. Among the most studied hardware-based approaches are multifocal and varifocal displays.

Multifocal displays address the problem by presenting imagery at more than one focal distance [59]. Instead of forcing accommodation to remain fixed at a single plane, these systems provide multiple focal planes so that virtual content can be displayed with more appropriate focus cues across depth. Prior work has shown that reducing the VAC in this way can improve user performance in 3D interaction tasks. For example, Batmaz et al. [44] compared multifocal and single-focal display conditions and found that reducing the conflict led to better virtual-hand pointing performance in immersive displays. These findings suggest that when focus cues are brought closer to the simulated depth of the target, users can interact more efficiently.

Varifocal displays follow a different strategy [59]. Rather than presenting several fixed focal planes, they dynamically adjust the focal distance of the image to better match the depth of the currently viewed content. In principle, this allows the system to more closely reproduce the natural coupling between vergence and accommodation during viewing. Recent work has shown that varifocal designs can reduce the impact of the VAC on 3D pointing performance relative to fixed-focal baselines [60]. This makes varifocal systems particularly promising for XR applications in which targets appear at a range of depths and accurate depth-guided interaction is required.

Although these approaches have shown clear benefits, they have not yet replaced single-focal stereo displays in current commercial HMDs [59]. Prior work in VR interaction has shown that commonly used head-mounted displays operate with fixed focal distances, meaning that accommodation remains tied to the display plane while vergence follows simulated depth [36, 56]. As a result, the VAC remains a practical limitation in current systems. For this reason, understanding how the VAC affects user performance in existing commercial hardware remains important, even as improved display technologies continue to be developed. These limitations are particularly relevant in the context of 3D interaction, where users must rely on accurate depth perception to guide movement and selection tasks.

2.3 3D Interaction and Selection in Virtual Reality

2.3.1 Fundamentals of 3D Target Selection

Performance in 3D target selection tasks remains lower than in comparable 2D pointing tasks [37, 61]. 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 measure of the speed–accuracy tradeoff, for targets that require movement in visual depth [36, 37, 58].

2.3.2 Effects of Depth on Interaction Performance

Depth itself is a contributor to pointing difficulty. Work on 3D pointing reports that movement direction and depth placement can affect movement time [61–63] and error rates [61, 62], and that movements along the depth axis can behave differently than lateral movements [61, 63]. Past work also observed the effect of depth for distal pointing, where distance-related factors can change effective measures and user strategies [62, 64].

2.4 Effects of VAC on 3D Interaction Performance

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 [44]. In a follow-up study, researchers showed that stereo deficiencies and target position relative to the focal plane significantly affected ray casting selection performance [45]. Past work also identified that biomechanical and ergonomic factors could modulate the detrimental effects of VAC during 3D selection [56].

However, these studies primarily assume that target depth remains constant during each selection movement. As a result, they do not account for situations in which depth changes continuously over time, which may introduce additional challenges for both perception and motor control.

2.5 Interaction Across Multiple and Dynamic Depths

2.5.1 Interaction with Moving Targets

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) [65]. Other work has modeled the relationship between arm movement and visual perception [66]. Cashion et al. [67] conducted a comprehensive study comparing four specific 3D selection techniques across scenarios with varying motion dynamics: Raycasting, SQUAD, Zoom, and Expand. Their evaluation revealed that while Raycasting is a fast baseline, it is highly problematic for dynamic targets because users must effectively "chase" the moving object around the screen to maintain cursor alignment. To mitigate this, they found that progressive refinement techniques like SQUAD and its variation, Expand, significantly improve accuracy for fast-moving targets. Haan et al. [68] 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.

2.5.2 Missing Perspective: VAC Under Dynamic Depth

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 these dynamic selection tasks.

2.6 Summary and Research Gap

The background reviewed in this chapter shows that depth perception in near-field interaction depends on the coordinated use of multiple visual cues, including stereopsis, motion parallax, vergence, and accommodation [31, 32]. Under natural viewing conditions, vergence and accommodation are tightly coupled, allowing the visual system to receive mutually consistent information about object distance [38, 52, 53]. In current single-focal XR systems, however, this coupling is disrupted because accommodation remains tied to the display focal plane while vergence follows the simulated depth of the stereo content [52, 56]. This mismatch gives rise to the vergence–accommodation conflict (VAC), which has been shown to affect visual comfort, oculomotor behavior, cognitive processing, and interaction performance [38, 40, 53].

Prior work in 3D interaction has also shown that depth is an important factor in pointing performance. Selection performance can vary depending on target placement in depth and the characteristics of the movement, including movements along the depth axis [62–64]. In addition, several studies have reported that stereo display limitations and VAC-related factors reduce 3D pointing performance, including movement time, accuracy, and throughput [36, 37, 44, 45, 56, 58]. More generally, prior work has also shown that modeling 3D pointing performance introduces additional complexity compared to planar tasks [61, 69, 70]. These findings establish that both visual depth and the quality of depth cues play a major role in interaction efficiency in XR.

At the same time, the existing literature leaves an important gap. Most prior studies examining the effect of the VAC on 3D interaction have focused on targets placed at fixed depths during each selection movement, even when different depth planes are used across trials or across conditions [44, 45, 56]. Similarly, work on dynamic target selection has mainly examined target motion, target characteristics, or interaction technique design without considering how a continuously changing

VAC may affect user performance [65–68]. As a result, prior findings do not fully explain user behavior in situations where a target moves toward or away from the user during the selection itself.

This limitation is important because continuously changing target depth creates a temporally evolving visual constraint. In such situations, vergence demand changes throughout the movement, while accommodation remains fixed at the focal distance of the display. Therefore, the magnitude of the VAC is not constant during the interaction, but changes over time. This means that results obtained from static-depth conditions cannot simply be assumed to generalize to tasks involving depth-varying targets. Understanding this case is particularly important for XR applications in which users interact with objects that approach, recede, or otherwise move in depth during task execution.

Motivated by this gap, this thesis investigates how continuously changing target depth affects 3D selection performance in a single-focal VR headset. Building on prior work on VAC, 3D pointing, and depth-dependent performance modeling, the thesis focuses specifically on the case in which the target depth changes during acquisition, creating a dynamic vergence–accommodation conflict. The following chapters present the experimental study, results, and analysis developed to examine this problem.

Chapter 3

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

The contents of this section is currently under double-blind review process at IEEE ISMAR 2026 conference: Zahra Rasoulifar, Rumeysa Turkmen, Mayra Donaji Barrera Machuca, Wolfgang Stuerzlinger, Anil Ufuk Batmaz. Dynamic Vergence–Accommodation Conflict: Effects on 3D Selection Performance When Target Depth Changes. Submitted to *Proceedings of the 25th IEEE International Symposium on Mixed and Augmented Reality (ISMAR).*

3.1 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 [71]. These advances have made it possible for industries like healthcare, the military, and the arts and entertainment 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 [19], table tennis [72], surgical skills training [18], and object assembly [20]. Another example is doctors using augmented reality (AR) HMDs to help guide them in surgeries [28]. 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) [31, 32]. Modern XR HMDs still struggle to simulate monocular and binocular cues [33, 34]. Particularly, current commercial HMDs suffer from the vergence-accommodation conflict (VAC), 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) [34, 35] that affects user performance in 3D selection tasks [36, 37]. Previous work has shown that the VAC produces a combination of increased cognitive and perceptual strain that impairs the user's ability to quickly and accurately perceive an object's spatial position [38–40], thereby affecting the speed and precision of interaction with the object [41–43].

Past work has investigated the effect of the VAC on user performance under different conditions, such as selection tasks occurring in peri-personal space [37, 43] or during distal pointing [45]. More recently, Bashar et al. [41] 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 target [49, 73], with past work identifying a high degree of variability between individuals [74] and a higher reliance on stereo depth cues than monocular cues [51]. In target selection using XR HMDs, 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 targets moving in depth. 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 VAC's magnitude changes over time. Such changes can negatively affect how users visually select targets, estimate depth, and coordinate eye–hand movements during selection. 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 to 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.

3.2 Previous Work

3.2.1 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 [38]. Related work has discussed visual comfort in binocular and 3D displays [55]. The VAC has also been connected to changes in visual system behavior (e.g., vergence and accommodation responses) when viewing stereoscopic imagery [53, 54] and to measurable costs in demanding cognitive tasks under induced conflict [40].

Commercial VR and AR HMDs like the Quest 3, Varjo-4, and Galaxy XR remain effectively single-focal systems due to the use of stereo displays at fixed focal distances as their underlying technology [34, 35]. 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 [52, 56].

Other work has compared different display types to determine whether they can overcome the VAC. Examples include multifocal and single-focal display conditions [44] and varifocal displays [60]. Results show that these technologies reduce the impact of the VAC on user performance. However, these approaches have not been widely adopted in current commercial HMDs.

3.2.2 VAC and 3D Interaction Performance

Stereo displays can improve 3D target selection, but performance in 3D remains lower than in comparable 2D pointing tasks [61]. 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 [36, 37, 58].

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 [44]. In a follow-up study, researchers showed that stereo deficiencies and target position relative to the focal plane significantly affected ray casting selection performance [45]. Past work also identified that biomechanical and ergonomic factors could increase the detrimental effects of VAC during 3D selection [56].

3.2.3 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 [61–63] and error rates [61, 62], and that movements along the depth axis can behave differently than lateral movements [61, 63]. Past work also observed the effect of depth for distal pointing, where distance-related factors can change effective measures and user strategies [62, 64].

Other studies used multiple depth planes or targets placed at different depths to study selection when visual depth changes. Examples include using a varying VAC condition, where target positions switch between the HMD's focal plane and locations away from the focal plane [45, 56]. These works have shown that the VAC affects the selection performance of the user. Yet, in all conditions, the position of the selected target remains constant during each individual 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) [65]. Other work has modeled the relationship between arm movement and visual perception [66]. Cashion et al. [67] conducted a comprehensive study comparing four specific 3D selection techniques across scenarios with varying motion dynamics: Raycasting, SQUAD, Zoom, and Expand. Their evaluation revealed that while Raycasting is a fast baseline, it is highly problematic for dynamic targets because users must effectively "chase" the moving object around the screen to maintain cursor alignment. To mitigate this, they found that progressive refinement techniques like SQUAD and its variation, Expand, significantly improve accuracy for fast-moving targets. Haan et al. [68] 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 these dynamic selection tasks.

3.2.4 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 [75, 76]. 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 tradeoff [77]. The ISO 9241-411 standard [78] defines a multi-directional selection task for evaluating pointing performance with Equation 2.

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

In Equation 2, effective throughput is calculated through the effective index of Difficulty divided by the movement time (MT). The ID_e is defined in Equation 3.

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

In Equation 3, the A_e represents the 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 [44]. Effective throughput and ISO 9241:411 multidirectional selection tasks have been widely adopted in HCI research [77] and for VR studies investigating VAC [56].

Different formulations of Fitts' law have been proposed for 3D interaction, as interaction techniques can involve different movement characteristics and control parameters [37, 41, 63, 69, 70]. For raycasting and other rotational pointing techniques, angular formulations are often used because the motor action is largely wrist rotation and the relevant geometry is better captured by angular separation and angular target width [64]. 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 [56].

3.2.5 Controlling Visual Angle and Perceived Size

Fitts' law shows that movement time is modeled as a function of movement amplitude and target width [75, 77]. 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 [61]. To address this, previous VAC studies scaled target size and target distance to maintain a constant visual angle across depth conditions, ensuring that targets were perceived with the same apparent size across depth planes. [43, 44, 56].

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 perceived target position and target size constant, one can vary vergence demand over time while avoiding additional size cues during target selection.

3.3 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 [38, 52, 56]. 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 [44, 45, 56]. 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 [49, 73]. 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 under static VAC conditions extend to this type of continuously changing, but perceptually subtle depth change. Based on the above motivation, this work is guided by the following main 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?*

3.4 Methodology

3.4.1 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 the local university research ethics board (protocol anonymized for review). The certification document is provided in Appendix A.2. Participation was voluntary and uncompensated, and no participants were excluded from the study.

3.4.2 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 1.33 m [41]. The virtual environment was developed using *Unity* version 2021.3.24f1 and displayed in the HMD via Meta Quest Link. A detailed visualization of the experimental setup and Dynamic VAC condition is provided in Appendix A.1.

3.4.3 Procedure

The user study was conducted on-site in an indoor laboratory on a 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 3.1.(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" [79], where mechanical input on the controller can unintentionally disturb pointing accuracy.

We used a variation of ISO 9241-411 multidirectional selection task [78] and adapted it to

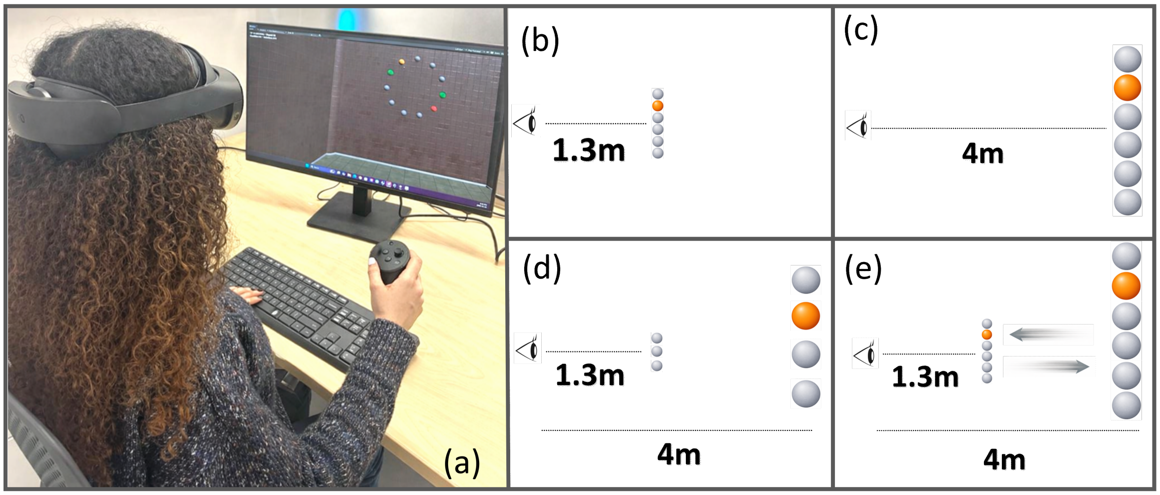


Figure 3.1: Overview of the experimental setup and Vergence–Accommodation Conflict (VAC) conditions. (a) A participant performing the ISO 9241-411 multidirectional selection task in VR while pointing with a controller and confirming selections using a keyboard. (b) *No VAC*: targets are located at the headset's focal plane (1.33 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 perceived size, causing the vergence–accommodation conflict to change over time.

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 3.1(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 3.1(b)).
- *Constant VAC*: Targets where placed at a far distance (Figure 3.1(c)). We used 4 m, as in previous work [45].
- *Varying VAC*: Half of the targets were placed on the focal plane and the other half at the far distance (Figure 3.1(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 time, taking 5 seconds to move from the near depth to the far depth and vice versa (Figure 3.1(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 and the effect of the continuously changing VAC could be isolated.

Target selection was performed using a small spherical pointer controlled by the VR controller. Although the pointer position was computed using an invisible ray, the ray itself was not rendered to avoid making depth changes visually noticeable, particularly in the Dynamic VAC condition. The pointer's apparent size was adjusted to remain visually consistent across different depths. 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 was used to reduce the possibility of revealing depth differences through head motion, particularly in the Varying VAC and Dynamic VAC conditions. At the beginning 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 the conditions 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. Training data were 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) [80], NASA TLX [81], 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.

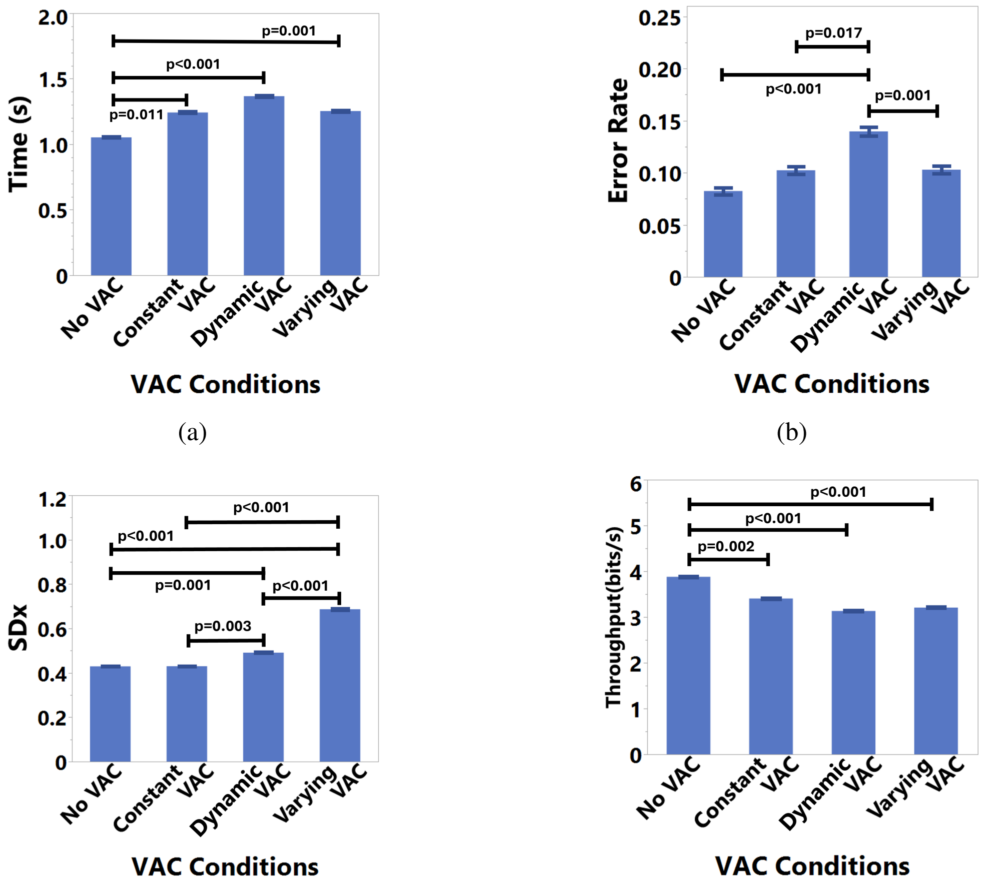
Participants were given short breaks between VAC conditions, during which they completed questionnaires related to the preceding condition before continuing the experiment.

3.4.4 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 [78] with 11 targets. To vary angular index of difficulty (ID_{α}), 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 [45]. Within each VAC condition and repetition, the nine ID configurations were presented in randomized order. Overall, each participant completed $4_{VAC} \times 3$ repetitions $\times 9_{ID_A} \times 11$ trials $= 1188$ selections.

3.5 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 [82, 83]. When this assumption was violated (even after log transformation), we applied the Aligned Rank Transform (ART) [84] 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 was analyzed with a Friedman's test and a Wilcoxon Signed Rank test for pairwise comparisons.



(a) (b)

(c) (d)

Figure 3.2: RM ANOVA results for (a) time, (b) error rate, (c) SD_x, and (d) throughput

3.5.1 Performance Results

Movement Time (MT). Movement time was significantly different for VAC conditions (Table 3.1). As shown in Fig. 3.2a, 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 3.1). As shown in Fig. 3.2b, 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.

Table 3.1: ANOVA results for VAC condition, index of difficulty (ID), and their interaction across performance measures. Significant effects ($p < 0.05$) are shown in bold.

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

SD_x. Angular endpoint variability differed significantly across VAC conditions (Table 3.1). As illustrated in Fig. 3.2c, 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.

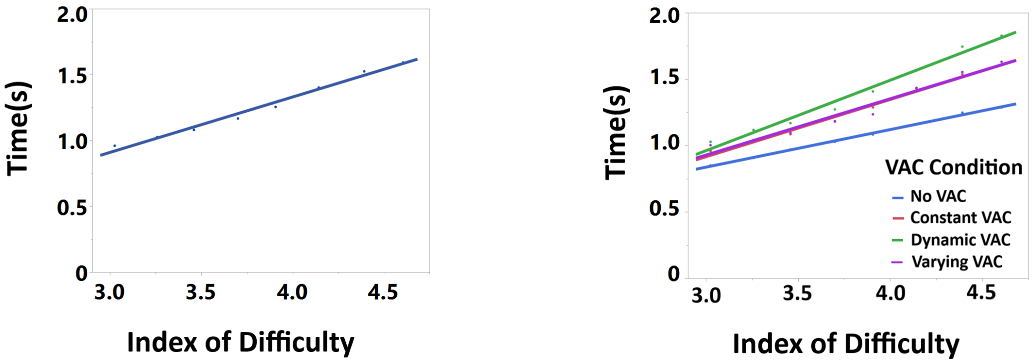
Throughput. Throughput differed significantly across VAC conditions (Table 3.1). As shown in Fig. 3.2d, the *No VAC* condition exhibited overall the highest throughput. Post hoc comparisons indicated that throughput in the *No VAC* condition was significantly higher than in the *Constant VAC* ($p = 0.002$), *Dynamic VAC* ($p < 0.001$), and *Varying VAC* ($p < 0.001$) conditions.

3.5.2 Fitts' Law Model Fit

Figure 3.3 shows the relationship between movement time and index of difficulty (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. 3.3a).

A separate analysis for each VAC condition showed that movement time increases as ID increases (Fig. 3.3b). 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.

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



(a) (b)

Figure 3.3: Fitts' law analysis for (a) the whole user study, and (b) each VAC condition.

scenario. Unlike the other experimental conditions, target depth in *Dynamic VAC* 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.

3.5.3 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 the index of difficulty (ID) and depth (Figure 3.4). 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:

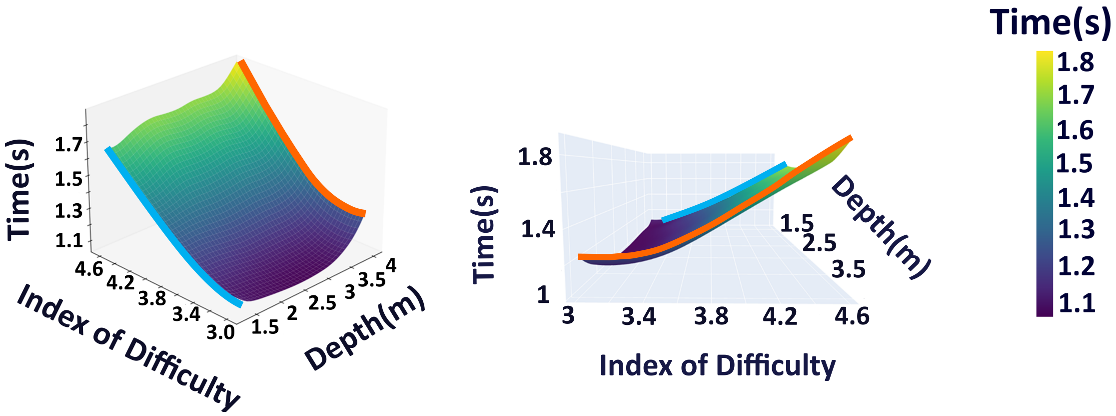


Figure 3.4: Smoothed empirical movement time surface for the Dynamic VAC condition as a function of index of difficulty (ID) and target depth. Color indicates movement time in seconds. The same surface is shown from two viewing angles (front and side) to better illustrate the relationship between task difficulty and depth.

the classical Fitts' law formulation [75, 76], a linear depth–interaction model, a diopter-based VAC model [41], 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 3.2: Comparison of candidate models for the Dynamic VAC condition. Lower AIC [4] and BIC [5] values indicate better model fit. Note that for nonlinear models in the last two rows, R^2 should *not* be used as a measure for comparison [6].

Model	**Formula**	**Coefficients**	R^2	**AIC**	**BIC**
1. Classical Fitts' Law [75, 76]	$\mathrm{MT} = a + b \cdot ID$	$a = 0.438, \quad b = 0.501$	0.63	−792.4	−784.9
2. Diopter-Based (VAC) [41]	$\mathrm{MT} = a + b \cdot ID + c \cdot \lvert ViD \rvert$	$a = 0.394, \quad b = 0.476$ $c = 0.853$	0.76	−893.7	−882.0
3. Linear Depth–Interaction	$\mathrm{MT} = a + b \cdot ID + c \cdot \Delta + d \cdot ID \cdot \Delta$	$a = 0.381, \quad b = 0.503$ $c = 0.712, \quad d = -0.321$	0.79	−921.4	−906.4
4. Power Depth Model	$\mathrm{MT} = a + b \cdot ID + c \cdot (\Delta + 1)^f$	$a = 0.368, \quad b = 0.498$ $c = 0.591, \quad f = 1.47$	0.81	−948.3	−933.3
5. Depth-Curvature (Proposed)	$\mathrm{MT} = a + b \cdot ID + c \cdot \Delta$ $+ d \cdot ID \cdot \Delta + e \cdot \Delta \cdot ID^2$	$a = -0.3617, \quad b = 0.4324$ $c = 0.7110, \quad d = -0.3933$ $e = 0.0560$	0.84	**−981.9**	**−964.8**

$\Delta = \max(D - 1.33,\ 0)$; focal plane fixed at 1.33 m. The proposed model is highlighted.

Table 3.2 summarizes the comparison of these models. To compare models with different functional forms and numbers of parameters, we used the Akaike Information Criterion (AIC) [4] and the Bayesian Information Criterion (BIC) [5]. Lower AIC and BIC values indicate a better model fit, while also penalizing unnecessary model complexity. These criteria are particularly useful when comparing nonlinear models, where R^2 alone may not provide a reliable basis for model comparison [6]. 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:

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

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 3.5 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 farther depths.

To make this behavior easier to interpret, Figure 3.6 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:

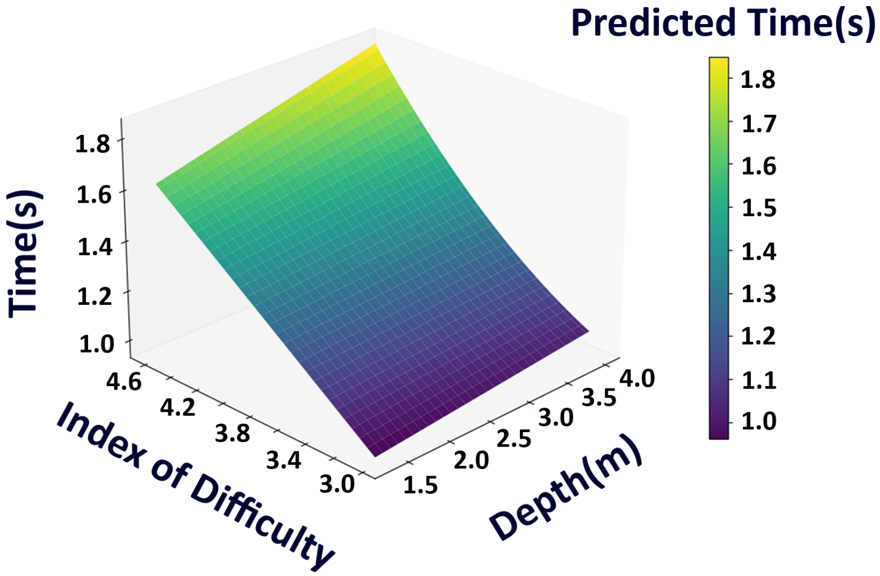


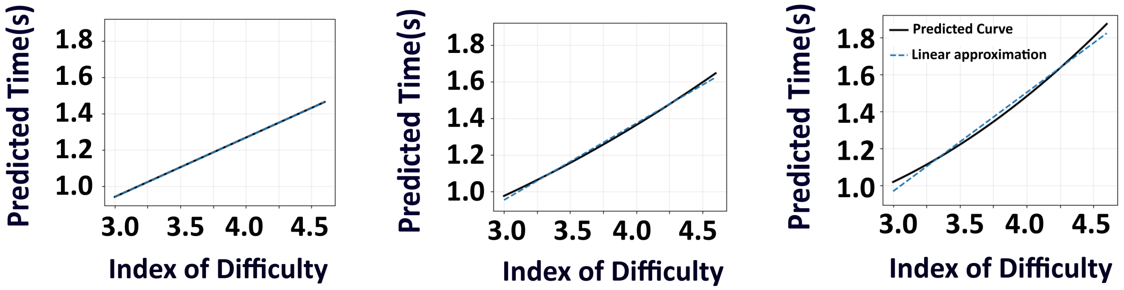
Figure 3.5: Movement time predicted by the fitted quadratic depth–curvature model as a function of index of difficulty (ID) and target depth in the Dynamic VAC condition. Depth is expressed relative to the headset focal plane (1.3 m). The model predicts that movement time increases with both task difficulty and depth, while the ID–time relationship becomes increasingly non-linear at farther depths.

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

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.

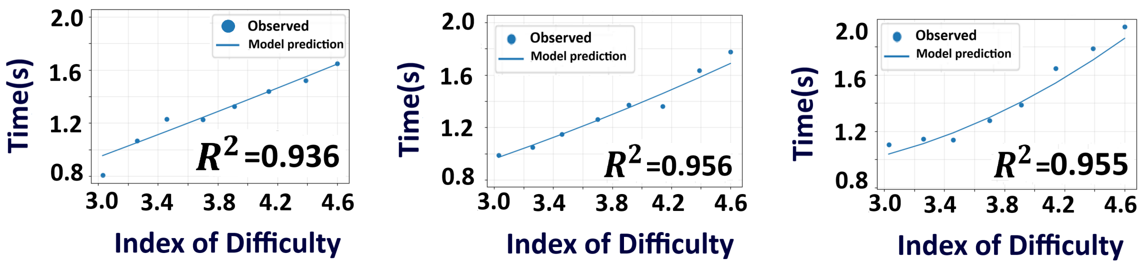
To further show how the proposed model captures the behavior of the Dynamic VAC data, we compared the model predictions with the observed movement times at some depths in the Dynamic VAC data (Fig. 3.7). The plots show the relationship between movement time and index of difficulty for depths within ±0.10 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.

As a consistency check, we evaluated how the model derived from the Dynamic VAC condition



(a) Depth = 1.3 m (b) Depth = 2.5 m (c) Depth = 4.0 m

Figure 3.6: 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 (ID) at a fixed depth. Near the focal plane, the fitted relationship is close to linear, while at farther depths it becomes increasingly curved.

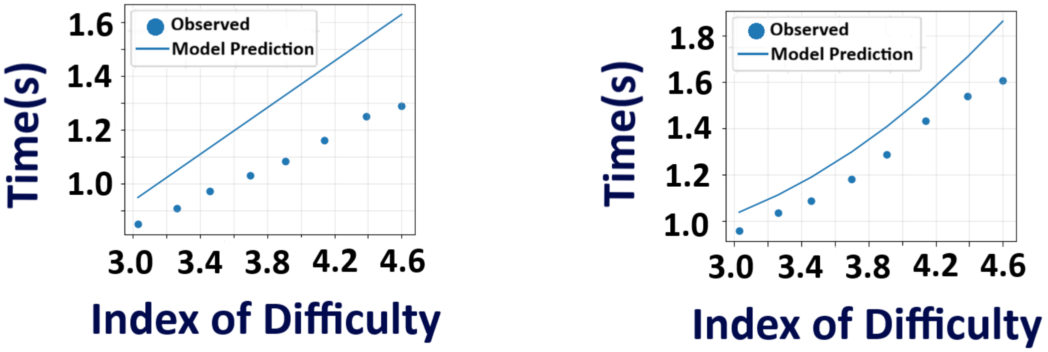


(a) Depth ≈ 1.5 m (b) Depth ≈ 2.0 m (c) Depth ≈ 4.0 m

Figure 3.7: 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.

relates to the static No VAC and Constant VAC conditions. To do 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 3.8 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. Movement time increases with task difficulty in a manner consistent with the model structure. However, 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 from Dynamic VAC overestimates movement time for stationary targets, but it captures the general relationship between task difficulty



(a) No VAC (targets at 1.3 m). (b) Constant VAC (targets at 4.0 m).

Figure 3.8: Comparison between observed movement times and predictions from the Dynamic VAC model for the static depth conditions. 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.

and performance across depth.

3.5.4 Subjective Measures

Cognitive Load Based on the NASA TLX scores, perceived workload differed significantly across VAC conditions (Fig. 3.9a). 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 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. 3.9. 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:** Temporal demand was significantly higher in both the *Dynamic VAC* ($p = 0.034$) and *Varying VAC* ($p = 0.002$) conditions 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* ($p = 0.042$).

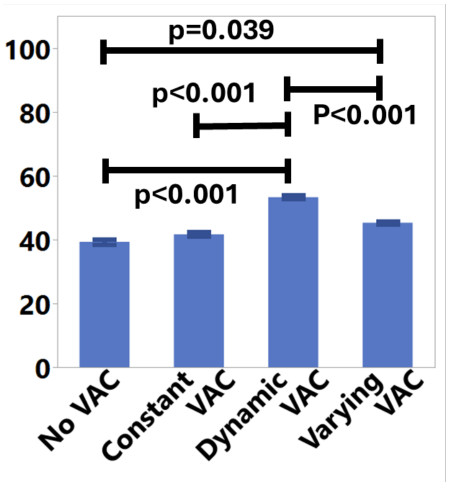
- **Frustration:** Frustration was significantly higher in the *Dynamic VAC* condition compared to *No VAC* ($p < 0.001$) and *Constant VAC* ($p = 0.015$). In addition, both *Constant VAC* ($p = 0.042$) and *Varying VAC* ($p = 0.003$) produced higher frustration than the *No VAC* condition.

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 3.3, 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 remained consistent and that differences observed in performance and workload measures are unlikely to be driven by interface usability differences.

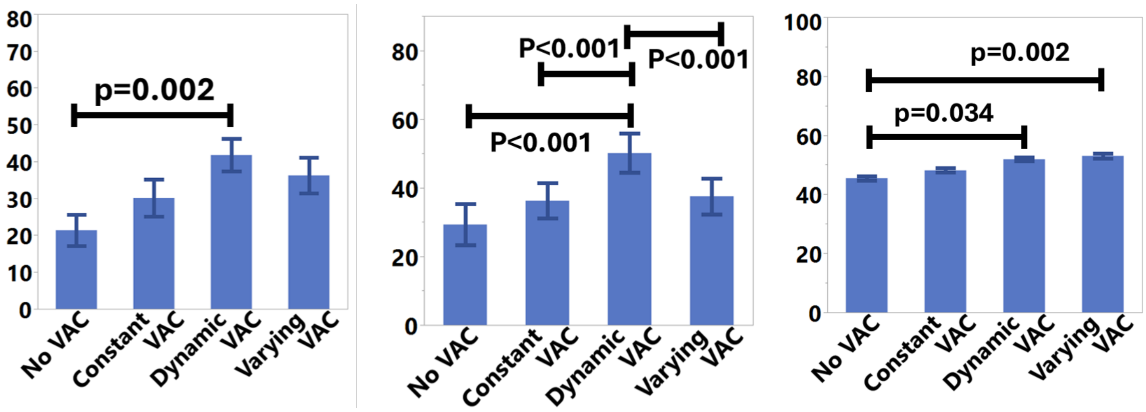
Table 3.3: SUS Results and Grades.

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 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



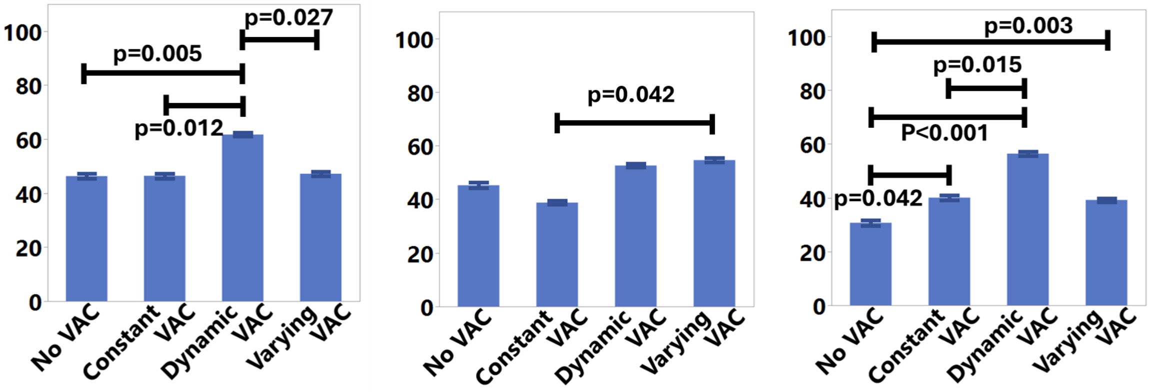
(a) Overall workload



(b) Mental demand

(c) Physical demand

(d) Temporal demand



(e) Performance

(f) Effort

(g) Frustration

Figure 3.9: 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.

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.

3.6 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, participants perceived the same target size even though the virtual depth of the targets changed.

3.6.1 Dynamic Target Selection

The motor performance results show that the VAC 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, indicating slower movement execution under the VAC. With the Dynamic VAC condition, we also observed 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 VAC, indicating an overall reduction in interaction efficiency.

Beyond confirming the performance cost of the VAC [56], 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, the visual constraint 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 should not simply be interpreted as 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 (including in the continuously changing depth dimension) and point to it, and adapt to the changing vergence demand. This additional instability in this condition causes an increased movement time and a 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 [49], 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.

3.6.2 Dynamic Vergence-Accommodation Conflict

Our findings are consistent with prior research showing that the VAC affects interaction performance in VR [36, 45, 56, 61]. Previous studies have 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.

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 itself is modulated by the magnitude of the VAC.

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 7124 data points for the Dynamic VAC condition. Since the target depth at selection was not constrained, data were distributed continuously across the depth range rather than accumulating at specific distances. A distribution analysis also confirmed that, with 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.

3.6.3 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 [72] and rhythm and action games [85]. 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**.

3.6.4 Limitations

The experiment was conducted using a single commercial VR HMD with a fixed focal plane. Since the magnitude of the VAC depends on the optical properties of the display, the observed effects may differ across headsets with different focal distances or optical designs. Second, the task 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, the Dynamic VAC condition involved controlled depth changes along the viewing direction while maintaining a constant visual angle, which isolates the effect of a changing vergence demand but does not capture all types of depth motion that may occur in practice. Changing the speed of the targets' movement might change the effects of VAC and the results here.

Future work could examine how 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 also investigate different depth motion profiles, speeds, and directions to determine how the rate of vergence change affects performance.

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.

3.7 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.

Chapter 4

Conclusion and Future Work

4.1 Conclusion

This thesis examined how continuously changing target depth affects 3D target selection in single-focal virtual reality head-mounted displays. While previous work on the vergence– accommodation conflict (VAC) has mainly studied stationary targets at fixed depths, this thesis focused on a different scenario: target selection when depth changes during the movement itself. In this situation, vergence demand does not remain stable throughout the selection, while accommodation stays fixed at the headset focal plane. This creates a Dynamic VAC condition that differs from previously studied static-depth conditions.

The results showed that the Dynamic VAC condition reduces interaction performance in terms of time, error rate, accuracy, and throughput, compared to stationary target conditions in which targets are positioned at the focal plane of the VR HMD. Moreover, the Dynamic VAC condition significantly increased number of errors and reduced the error compares to stationary target conditions. These findings indicate that the effect of depth in XR interaction should not be treated as constant during target acquisition. When target depth changes continuously, the associated visual constraints also evolve throughout the movement, making the task more demanding for both perceptual processing and motor control.

Our further analysis based on Fitts' law revealed consistent findings, indicating that the existing formulation does not adequately capture performance changes when targets move along the depth axis. When we applied both the traditional and recently proposed model formulations, the results showed relatively low goodness-of-fit. However, when movement time was analyzed jointly with task difficulty and target depth, the data exhibited a quadratic trend. This observation suggests that user performance in such dynamic depth conditions cannot be sufficiently explained by a linear formulation. Based on these findings, we proposed a novel Fitts' law model that incorporates both task difficulty and depth variation. The proposed model demonstrated improved fit compared to existing formulations at different target depths.

Overall, this thesis contributes to a better understanding of how depth-dependent visual constraints affect 3D interaction in VR. These findings help extend the study of VAC from fixed-depth selection toward more realistic interactive scenarios in which objects move during the task.

4.2 Future Work

One potential research direction is to investigate Dynamic VAC across different interaction techniques, such as virtual-hand interaction, gaze-based selection, and multi-modal approaches. These techniques rely on distinct balances between motor execution, visual feedback, and control-display mappings. For example, recent work on depth-adaptive gaze selection with stationary targets [16] could be extended to evaluate performance under continuously changing target depth. Therefore, the performance effects observed in this thesis may not transfer uniformly across interaction modalities. Systematic comparisons could help identify how Dynamic VAC interacts with modality-specific constraints, such as hand-tracking precision, dwell-based selection dynamics, or tool-mediated control.

Another direction is to examine target motion under varying temporal and spatial characteristics. In this thesis, depth variation followed a controlled and continuous trajectory within predefined motion ranges (1.33m—4m). Future studies can investigate whether similar performance patterns emerge under different motion speeds, depth planes, acceleration profiles, and target behaviors (e.g., sudden depth changes or unpredictable motion). Such analyses would help identify which properties of depth motion contribute most strongly to increased task difficulty and performance costs.

Dynamic VAC should also be evaluated in ecologically valid XR scenarios. Real-world XR interactions often involve objects that move toward or away from the user during manipulation, training, or gameplay. Studying Dynamic VAC in applied contexts—such as skill training simulations, game-like environments, or collaborative tasks—would provide insight into how the controlled experimental findings reported here translate to practical XR use cases.

Future work can also explore whether advanced display technologies mitigate the observed performance costs. Since the primary source of conflict in this thesis arise from the fixed focal plane of current single-focal HMDs, it would be valuable to replicate the same experimental framework using multifocal or varifocal displays. Such comparisons could clarify whether reducing the vergence–accommodation mismatch also alleviates the additional demands introduced by continuously changing target depth.

Appendix A

Appendix

A.1 Experimental Setup

This section presents additional visual details of the experimental setup and illustrates the Dynamic VAC condition across time.

A.2 Ethics Approval

The following document shows the certification of ethical acceptability for this study.

A.3 Accepted & Submitted Papers

- **Zahra Rasoulifar**, Rumeysa Turkmen, Mayra Donaji Barrera Machuca, Wolfgang Stuerzlinger, Anil Ufuk Batmaz. Dynamic Vergence–Accommodation Conflict: Effects on 3D Selection Performance When Target Depth Changes. Submitted to *Proceedings of the 25th IEEE International Symposium on Mixed and Augmented Reality (IEEE ISMAR 2026).*
- Rumeysa Turkmen, **Zahra Rasoulifar**, Anil Ufuk Batmaz. VoiceRay: 3D Object Selection Technique for Occluded and Dense Environments in Virtual Reality. In *Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems, CHI'26, pages 1–14, New York, NY, USA, 2026. ACM.*

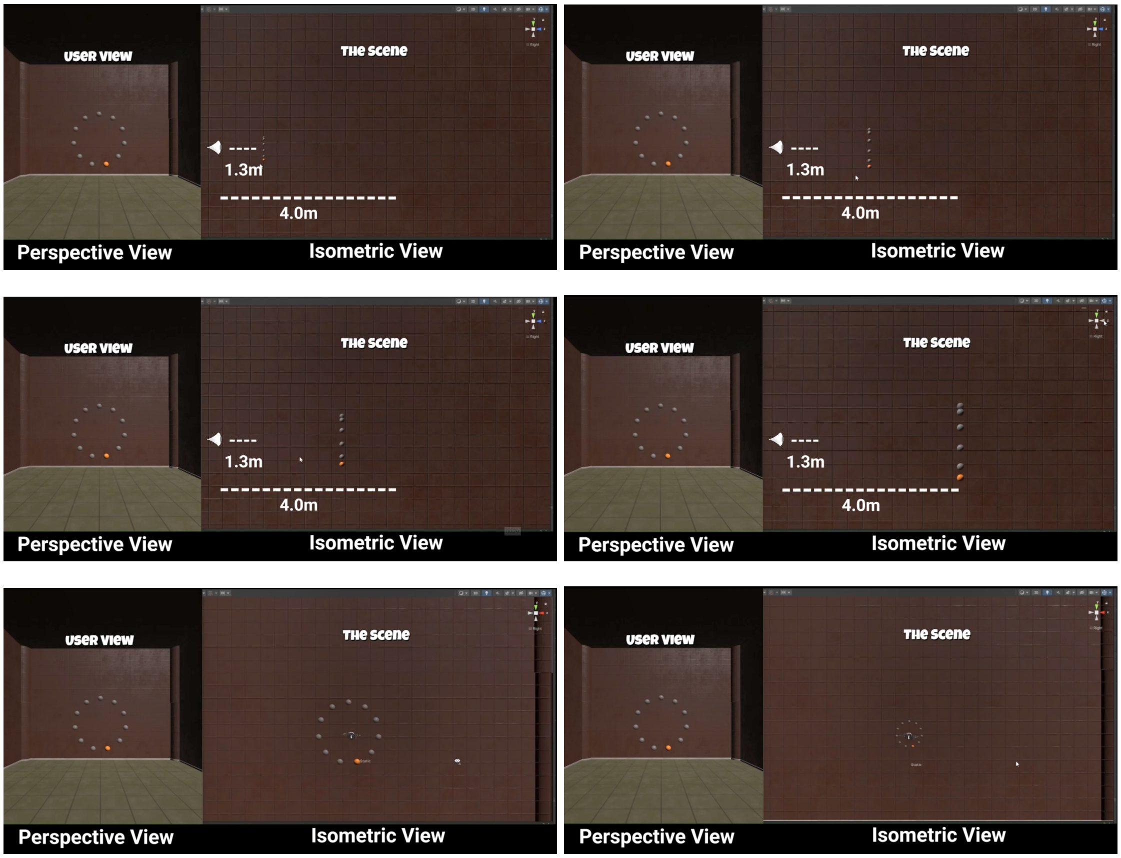


Figure A.1: Additional illustrations of the experimental setup and Dynamic VAC condition. Each frame presents the participant's perspective view (left) alongside the corresponding isometric scene view (right). Across the six frames, the highlighted target is shown at different positions along the depth axis (1.3 m to 4.0 m), while the circular arrangement of targets remains visually stable from the user's perspective. These examples further illustrate how dynamic depth changes were introduced during the selection task without altering the apparent layout of targets.

UNIVERSITÉ
Concordia
UNIVERSITY

CERTIFICATION OF ETHICAL ACCEPTABILITY
FOR RESEARCH INVOLVING HUMAN SUBJECTS

Name of Applicant:	Zahra Rasoulifar
Department:	Gina Cody School of Engineering and Computer Science\Computer Science and Software Engineering
Agency:	N/A
Title of Project:	Investigating Target Depth Through Object Selection in Virtual Environments
Certification Number:	30022736

Valid From: November 20, 2025 To: November 19, 2026

The members of the University Human Research Ethics Committee have examined the application for a grant to support the above-named project, and consider the experimental procedures, as outlined by the applicant, to be acceptable on ethical grounds for research involving human subjects.

Richard DeMont

Dr. Richard DeMont, Chair, University Human Research Ethics Committee

Figure A.2: Certification of ethical acceptability for research involving human subjects.

- Amal Hatira, **Zahra Rasoulifar**, Mucahit Gemici, Vrushank Phadnis, Anil Ufuk Batmaz. Effects of False Positive and False Negative Early Warnings for Hand Tracking Failures on User Trust and Confidence in Virtual Reality. Submitted to *Proceedings of the 25th IEEE International Symposium on Mixed and Augmented Reality (IEEE ISMAR 2026).*

- Rumeysa Turkmen, **Zahra Rasoulifar**, Aunnoy K Mutasim, Anil Ufuk Batmaz. Investigating Depth-Based Target Expansion for Gaze Selection with Dwell Activation in Virtual Environments. Submitted to *Proceedings of the 25th IEEE International Symposium on Mixed and Augmented Reality (IEEE ISMAR 2026).*

Bibliography

[1] Paul Milgram, Haruo Takemura, Akira Utsumi, and Fumio Kishino. Augmented reality: A class of displays on the reality-virtuality continuum. In *Telemanipulator and telepresence technologies*, volume 2351, pages 282–292. Spie, 1995.

[2] Interaction Design Foundation. VR headsets, 2023. URL `https://ixdf.org/literature/topics/vr-headsets`. Accessed: 2026-03-21.

[3] Meta Platforms, Inc. Meta quest pro: Price, release date, and specs, 2022. URL `https://www.meta.com/blog/meta-quest-pro-price-release-date-specs/`. Accessed: 2026-03-21.

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

[6] Andrej-Nikolai Spiess and Natalie Neumeyer. An evaluation of R^2 as an inadequate measure for nonlinear models in pharmacological and biochemical research: A Monte Carlo approach. *BMC Pharmacology*, 10(1):6, 2010.

[7] Joseph J. LaViola Jr., Ernst Kruijff, Ryan P. McMahan, Doug Bowman, and Ivan P. Poupyrev. *3D User Interfaces: Theory and Practice*. Addison-Wesley Professional, 2017.

[8] Hyejin Kim, Sukwon Lee, and Changgu Kang. From controllers to multimodal input: A chronological review of xr interaction across device generations. *Sensors*, 26(1):196, 2025.

[9] Rumeysa Turkmen, Zahra Rasoulifar, and Anil Ufuk Batmaz. VoiceRay: 3D object selection technique for occluded and dense environments in virtual reality. In *Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems*, CHI '26, pages 1–14, New York, NY, USA, 2026. ACM.

[10] Michail Pavlou, Dimitrios Laskos, Evangelia I Zacharaki, Konstantinos Risvas, and Konstantinos Moustakas. Xrsise: An xr training system for interactive simulation and ergonomics assessment. *Frontiers in Virtual Reality*, 2:646415, 2021.

[11] Anders Berglund et al. Design for extended reality (dfxr)–exploring engineering and product design education in xr. In *DS 123: Proceedings of the International Conference on Engineering and Product Design Education (E&PDE 2023)*, 2023.

[12] Shah Mahsoom Ali, Satyabrata Aich, Ali Athar, and Hee-Cheol Kim. Medical education, training and treatment using xr in healthcare. In *2023 25th international conference on advanced communication technology (ICACT)*, pages 388–393. IEEE, 2023.

[13] Ahmed Alnagrat, Rizalafande Che Ismail, Syed Zulkarnain Syed Idrus, and Rawad Mansour Abdulhafith Alfaqi. A review of extended reality (xr) technologies in the future of human education: Current trend and future opportunity. *Journal of Human Centered Technology*, 1 (2):81–96, 2022.

[14] Yongjae Lee and Byounghyun Yoo. Xr collaboration beyond virtual reality: work in the real world. *Journal of Computational Design and Engineering*, 8(2):756–772, 2021.

[15] Jannick P Rolland and Hong Hua. Head-mounted display systems. *Encyclopedia of optical engineering*, 2:1–14, 2005.

[16] Rumeysa Turkmen, Zahra Rasoulifar, Aunnoy K. Mutasim, and Anil Ufuk Batmaz. Investigating depth-based target expansion for gaze selection with dwell activation in virtual environments, 2026. Submitted to Proceedings of the 25th IEEE International Symposium on Mixed and Augmented Reality (ISMAR '26).

[17] Steven M. LaValle. *Virtual Reality*. Cambridge University Press, Cambridge, 2017. URL http://lavalle.pl/vr/.

[18] Neal E. Seymour, Anthony G. Gallagher, Sanziana A. Roman, Michael K. O'Brien, Vipin K. Bansal, Dana K. Andersen, and Richard M. Satava. Virtual reality training improves operating room performance: Results of a randomized, double-blinded study. *Annals of Surgery*, 236 (4):458–464, 2002.

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

[20] Amal Hatira, Zahra Rasoulifar, Mucahit Gemici, Vrushank Phadnis, and Anil Ufuk Batmaz. Effects of false positive and false negative early warnings for hand tracking failures on user trust and confidence in virtual reality, 2026. Submitted to Proceedings of the 25th IEEE International Symposium on Mixed and Augmented Reality (ISMAR '26).

[21] Leif P. Berg and Judy M. Vance. Industry use of virtual reality in product design and manufacturing: A survey. *Virtual Reality*, 21(1):1–17, 2017. doi: 10.1007/s10055-016-0293-9.

[22] Kate E. Laver, Belinda Lange, Stacey George, Judith E. Deutsch, Gustavo Saposnik, and Maria Crotty. Virtual reality for stroke rehabilitation. *Cochrane Database of Systematic Reviews*, (11):CD008349, 2017. doi: 10.1002/14651858.CD008349.pub4.

[23] Federica Pallavicini, Alessandro Pepe, and Maria Eleonora Minissi. Gaming in virtual reality: What changes in terms of usability, emotional response and sense of presence compared to non-immersive video games? *Simulation & Gaming*, 50(2):136–159, 2019. doi: 10.1177/1046878119831420.

[24] Ronald T. Azuma. A survey of augmented reality. *Presence: Teleoperators and Virtual Environments*, 6(4):355–385, 1997. doi: 10.1162/pres.1997.6.4.355.

[25] Ronald T. Azuma, Yohan Baillot, Reinhold Behringer, Steven K. Feiner, Simon Julier, and

Blair MacIntyre. Recent advances in augmented reality. *IEEE Computer Graphics and Applications*, 21(6):34–47, 2001. doi: 10.1109/38.963459.

[26] Ana Javornik. Augmented reality: Research agenda for studying the impact of its media characteristics on consumer behaviour. *Journal of Retailing and Consumer Services*, 30:252–261, 2016. doi: 10.1016/j.jretconser.2016.02.004.

[27] Jens Grubert, Tobias Langlotz, Stefanie Zollmann, and Holger Regenbrecht. Towards pervasive augmented reality: Context-awareness in augmented reality. *IEEE Transactions on Visualization and Computer Graphics*, 23(6):1706–1724, 2017. doi: 10.1109/TVCG.2016.2543720.

[28] Sulaman Durrani, Chiduziem Onyedimma, Ryan Jarrah, Atiq Bhatti, Karim Rizwan Nathani, Archis R. Bhandarkar, William Mualem, Abdul Karim Ghaith, Cameron Zamanian, Giorgos D. Michalopoulos, A. Yohan Alexander, Walter Jean, and Mohamad Bydon. The virtual vision of neurosurgery: How augmented reality and virtual reality are transforming the neurosurgical operating room. *World Neurosurgery*, 168:190–201, 2022. ISSN 1878-8750. doi: 10.1016/j.wneu.2022.10.002. URL `https://www.sciencedirect.com/science/article/pii/S1878875022014073`.

[29] Vasudev Bhaskaran and Upal Mahbub. Immersive user experiences: trends and challenges of using xr technologies. *Computer Vision*, pages 260–278, 2024.

[30] Becky Spittle, Maite Frutos-Pascual, Chris Creed, and Ian Williams. A review of interaction techniques for immersive environments. *IEEE Transactions on Visualization and Computer Graphics*, 29(9):3900–3921, 2022.

[31] James E. Cutting and Peter 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*, pages 69–117. Academic Press, San Diego, CA, USA, 1995. doi: 10.1016/B978-012240530-3/50005-5.

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

[33] Claudia Armbüster, Marc Wolter, Torsten Kuhlen, Will Spijkers, and Bruno Fimm. Depth perception in virtual reality: Distance estimations in peri- and extrapersonal space. *CyberPsychology & Behavior*, 11(1):9–15, 2008. doi: 10.1089/cpb.2007.9935.

[34] Haley Adams, Jeanine Stefanucci, Sarah Creem-Regehr, and Bobby 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)*, pages 792–801, 2022. doi: 10.1109/VR51125.2022.00101.

[35] Kevin Pfeil, Sina Masnadi, Jacob Belga, Jose-Valentin T. Sera-Josef, and Joseph LaViola. Distance perception with a video see-through head-mounted display. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pages 1–9, 2021. doi: 10.1145/3411764.3445223.

[36] Anil Ufuk Batmaz, Mayra Donaji Barrera Machuca, Duc Minh Pham, and Wolfgang Stuerzlinger. Do head-mounted display stereo deficiencies affect 3D pointing tasks in AR and VR? In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pages 585–592, 2019. doi: 10.1109/VR.2019.8797975.

[37] Mayra Donaji Barrera Machuca and Wolfgang 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, pages 1–14, New York, NY, USA, 2019. ACM. ISBN 9781450359702. doi: 10.1145/3290605.3300437. URL `https://doi.org/10.1145/3290605.3300437`.

[38] David M. Hoffman, Ahna R. Girshick, Kurt Akeley, and Martin S. Banks. Vergence–accommodation conflicts hinder visual performance and cause visual fatigue. *Journal of Vision*, 8(3):33, 03 2008. doi: 10.1167/8.3.33. URL `https://doi.org/10.1167/8.3.33`.

[39] Cyril Vienne, Stéphane Masfrand, Christophe Bourdin, and Jean-Louis 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.

[40] François Daniel and Zoï Kapoula. Induced vergence-accommodation conflict reduces cognitive performance in the Stroop test. *Scientific Reports*, 9(1):1247, February 2019. doi: 10.1038/s41598-018-37814-z.

[41] Mohammad Raihanul Bashar, Mayra Donaji Barrera Machuca, Wolfgang Stuerzlinger, and Anil Ufuk Batmaz. The effect of visual depth on the vergence–accommodation conflict on 3D selection performance within virtual reality headsets. *The Visual Computer*, 41(12):9645–9661, 2025. ISSN 1432-2315. doi: 10.1007/s00371-025-03990-x. URL `https://doi.org/10.1007/s00371-025-03990-x`.

[42] Xiaoye Michael Wang, Daniel Southwick, Ian Robinson, Michael Nitsche, Gabby Resch, Ali Mazalek, and Timothy N. Welsh. The geometry of the vergence-accommodation conflict in mixed reality systems. *Virtual Reality*, 28(2):95, 2024.

[43] Anil Ufuk Batmaz, Moaaz Hudhud Mughrabi, Mine Sarac, Mayra Donaji Barrera Machuca, and Wolfgang Stuerzlinger. Measuring the effect of stereo deficiencies on peripersonal space pointing. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, VR '23, Mar 2023. doi: 10.1109/VR55154.2023.00063.

[44] Anil Ufuk Batmaz, Mayra Donaji Barrera Machuca, Junwei Sun, and Wolfgang 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, New York, NY, USA, 2022. ACM. ISBN 9781450391573. doi: 10.1145/3491102.3502067. URL `https://doi.org/10.1145/3491102.3502067`.

[45] Anil Ufuk Batmaz, Moaaz Hudhud Mughrabi, Mayra Donaji Barrera Machuca, and Wolfgang Stuerzlinger. Effect of stereo deficiencies on virtual distal pointing. In *Proceedings of the 28th ACM Symposium on Virtual Reality Software and Technology*, VRST '22, New York, NY, USA, 2022. ACM. ISBN 9781450398893. doi: 10.1145/3562939.3565621. URL `https://doi.org/10.1145/3562939.3565621`.

[46] Zahra Rasoulifar, Rumeysa Turkmen, Mayra Donaji Barrera Machuca, Wolfgang Stuerzlinger, and Anil Ufuk Batmaz Batmaz. Dynamic vergence–accommodation conflict: Effects on 3D

selection performance when target depth changes, 2026. Submitted to Proceedings of the 25th IEEE International Symposium on Mixed and Augmented Reality (ISMAR '26).

[47] David R. Melmoth and Simon Grant. Advantages of binocular vision for the control of reaching and grasping. *Experimental Brain Research*, 171(3):371–388, 2006. doi: 10.1007/s00221-005-0273-x.

[48] Ian P. Howard and Brian J. Rogers. *Seeing in Depth, Vol. 2: Depth Perception*. I Porteous, Toronto, 2002.

[49] Harold T. Nefs and Julie 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.

[50] W. Neil Charman. The eye in focus: Accommodation and presbyopia. *Clinical and Experimental Optometry*, 91(3):207–225, 2008. doi: 10.1111/j.1444-0938.2008.00256.x.

[51] Philip B. Kruger, Karan R. Aggarwala, Sharon 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.

[52] Martin S. Banks, Joohwan Kim, and Takashi Shibata. Insight into vergence-accommodation mismatch. In *Head- and Helmet-Mounted Displays XVIII: Design and Applications*, volume 8735 of *Proc. SPIE*, page 873509, 2013. doi: 10.1117/12.2019866.

[53] Cyril Vienne, Laurent Sorin, Laurent Blondé, Quan Huynh-Thu, and Pascal Mamassian. Effect of the accommodation-vergence conflict on vergence eye movements. *Vision Research*, 100: 124–133, 2014. ISSN 0042-6989. doi: 10.1016/j.visres.2014.04.017. URL `https://www.sciencedirect.com/science/article/pii/S0042698914001126`.

[54] Tetsuya Fukushima, Masahito Torii, Kazuhiko Ukai, James S. Wolffsohn, and Bernard Gilmartin. The relationship between CA/C ratio and individual differences in dynamic accommodative responses while viewing stereoscopic images. *Journal of Vision*, 9(13):21, 12 2009. doi: 10.1167/9.13.21. URL `https://doi.org/10.1167/9.13.21`.

[55] Frank L. Kooi and Alexander Toet. Visual comfort of binocular and 3D displays. *Displays*, 25(2):99–108, 2004. ISSN 0141-9382. doi: 10.1016/j.displa.2004.07.004. URL https://www.sciencedirect.com/science/article/pii/S0141938204000502.

[56] Anil Ufuk Batmaz, Rumeysa Turkmen, Mine Sarac, Mayra Donaji Barrera Machuca, and Wolfgang 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, New York, NY, USA, 2023. ACM. ISBN 9798400703287. doi: 10.1145/3611659.3615686. URL https://doi.org/10.1145/3611659.3615686.

[57] Paul Lubos, Gerd Bruder, and Frank Steinicke. Analysis of direct selection in head-mounted display environments. In *IEEE Symposium on 3D User Interfaces*, pages 11–18. IEEE, 2014. doi: 10.1109/3DUI.2014.6798834.

[58] Mayra Donaji Barrera Machuca and Wolfgang 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, pages 1–6, New York, NY, USA, 2018. ACM. ISBN 9781450356213. doi: 10.1145/3170427.3188540. URL https://doi.org/10.1145/3170427.3188540.

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

[60] Xiaodan Hu, Monica Perusquía-Hernández, Mayra Donaji Barrera Machuca, Anil Ufuk Batmaz, Yan Zhang, Wolfgang Stuerzlinger, and Kiyoshi Kiyokawa. Varifocal displays reduce the impact of the vergence-accommodation conflict on 3D pointing performance in augmented reality systems, 2026. URL https://arxiv.org/abs/2602.05129.

[61] Robert J. Teather and Wolfgang Stuerzlinger. Pointing at 3D targets in a stereo head-tracked virtual environment. In *IEEE Symposium on 3D User Interfaces*, 3DUI '11, pages 87–94, USA, 2011. IEEE Computer Society. ISBN 9781457700637.

[62] Izabelle Janzen, Vasanth K. Rajendran, and Kellogg 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, pages 188–199, New York, NY, USA, 2016. ACM. ISBN 9781450333627. doi: 10.1145/2858036.2858244. URL `https://doi.org/10.1145/2858036.2858244`.

[63] Atsuo Murata and Hirokazu Iwase. Extending Fitts' law to a three-dimensional pointing task. *Human Movement Science*, 20(6):791–805, 2001. ISSN 0167-9457. doi: 10.1016/S0167-9457(01)00058-6. URL `https://www.sciencedirect.com/science/article/pii/S0167945701000586`.

[64] Regis Kopper, Doug A. Bowman, Mara G. Silva, and Ryan P. McMahan. A human motor behavior model for distal pointing tasks. *International Journal of Human-Computer Studies*, 68(10):603–615, 2010. ISSN 1071-5819. doi: 10.1016/j.ijhcs.2010.05.001. URL `https://www.sciencedirect.com/science/article/pii/S1071581910000637`.

[65] Jin Huang, John J. Dudley, Stephen Uzor, Dong Wu, Per Ola Kristensson, and Feng Tian. Understanding user performance of acquiring targets with motion-in-depth in virtual reality. *International Journal of Human-Computer Studies*, 163:102817, 2022. ISSN 1071-5819. doi: 10.1016/j.ijhcs.2022.102817. URL `https://www.sciencedirect.com/science/article/pii/S1071581922000465`.

[66] Yawen Zheng, Jin Huang, Hao Zhang, Yulong Bian, Juan Liu, Chenglei Yang, Feng Tian, and Xiangxu Meng. 3D ternary-Gaussian model: Modeling pointing uncertainty of 3D moving target selection in virtual reality. *International Journal of Human-Computer Studies*, 198: 103454, 2025. ISSN 1071-5819. doi: 10.1016/j.ijhcs.2025.103454. URL `https://www.sciencedirect.com/science/article/pii/S1071581925000114`.

[67] Jeffrey Cashion, Chadwick Wingrave, and Joseph 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.

[68] Gerwin de Haan, Michal Koutek, and Frits H. Post. IntenSelect: Using dynamic object rating for assisting 3D object selection. In *Proceedings of the 11th Eurographics Conference on Virtual Environments*, EGVE '05, pages 201–209, Goslar, Germany, 2005. Eurographics Association. ISBN 3905673215.

[69] Yeonjoo Cha and Rohae Myung. Extended Fitts' law for 3D pointing tasks using 3D target arrangements. *International Journal of Industrial Ergonomics*, 43(4):350–355, 2013. ISSN 0169-8141. doi: 10.1016/j.ergon.2013.05.005. URL `https://www.sciencedirect.com/science/article/pii/S0169814113000723`.

[70] Logan D. Clark, Aakash B. Bhagat, and Sara 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. ISSN 1071-5819. doi: 10.1016/j.ijhcs.2020.102413. URL `https://www.sciencedirect.com/science/article/pii/S1071581920300173`.

[71] Vladislav Angelov, Emiliyan Petkov, Georgi Shipkovenski, and Teodor Kalushkov. Modern virtual reality headsets. In *International Congress on Human-Computer Interaction, Optimization and Robotic Applications*, pages 1–5, 2020. doi: 10.1109/HORA49412.2020.9152604.

[72] Eren Karatas, Kissinger Sunday, Sude Erva Apak, Yiwei Li, Junwei Sun, Anil Ufuk Batmaz, and Mayra Donaji Barrera Machuca. "I 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, New York, NY, USA, 2023. ACM. ISBN 9798400702815. doi: 10.1145/3607822.3614539. URL `https://doi.org/10.1145/3607822.3614539`.

[73] David 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.

[74] Harold T. Nefs, Louise O'Hare, and Julie M. Harris. Two independent mechanisms for motion-in-depth perception: Evidence from individual differences. *Frontiers in Psychology*, 1, 2010.

ISSN 1664-1078. doi: 10.3389/fpsyg.2010.00155. URL https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2010.00155.

[75] Paul 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.

[76] I. Scott 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.

[77] R. William Soukoreff and I. Scott 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. ISSN 1071-5819. doi: 10.1016/j.ijhcs.2004.09.001. URL https://www.sciencedirect.com/science/article/pii/S1071581904001016.

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

[79] Doug Bowman, Chadwick Wingrave, Joshua Campbell, and Vinh Ly. Using pinch gloves for both natural and abstract interaction techniques in virtual environments. In *HCI International*, 2001.

[80] John Brooke. SUS: A 'quick and dirty' usability scale. In *Usability Evaluation in Industry*. CRC Press, London, 1996.

[81] Sandra 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.

[82] Joseph F. Hair Jr., William C. Black, Barry J. Babin, and Rolph E. Anderson. Multivariate data analysis, 2014.

[83] Paul Mallery and Darren George. SPSS for windows step by step: A simple guide and reference. *Allyn & Bacon*, 2003.

[84] Jacob O. Wobbrock, Leah Findlater, Darren Gergle, and James J. Higgins. The aligned rank transform for nonparametric factorial analyses using only ANOVA procedures. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, CHI '11, pages 143–146, New York, NY, USA, 2011. ACM. ISBN 978-1-4503-0228-9. doi: 10.1145/1978942.1978963. URL `http://doi.acm.org/10.1145/1978942.1978963`.

[85] Beat Games. Beat saber, 2019. URL `https://beatsaber.com/`. Virtual reality rhythm game.