

Looking Around by Looking Around: Omnidirectional Gaze-based VR Viewport Control

Hock Siang Lee , Jinghui Hu , Florian Weidner , Haopeng Wang , and Hans Gellersen 

Abstract—Traditional VR viewport control primarily relies on head and torso movement, which can be effortful and limiting in both constrained and extended-use settings. We introduce Looking Around by Looking Around (LALA), a gaze-based VR pitch-and-yaw viewport control technique designed for natural and effortless omnidirectional exploration via eye movements, without requiring or obstructing movement of the head, hand, or body, offering a low-effort and highly accessible interaction method. Because gaze is primarily used for perception and exhibits oculomotor and perceptual asymmetries, using it directly for control is difficult. To address this, we designed an asymmetric omnidirectional control profile for the eye, then built on it to exploit tendencies for eyes to stay within comfortable regions for viewport control. We evaluated LALA in a user study (N=18) featuring two contrasting tasks: alignment towards known directions and open-ended visual search towards unknown directions. LALA was strongly preferred over the traditional baseline, achieving competitive performance while enabling fully hands-free interaction with minimal physical movement.

Index Terms—Eye tracking, gaze interaction, eye-head coordination, viewport control, user study, gaze-based interaction, virtual reality.

1 INTRODUCTION

Virtual reality (VR) enables users to explore immersive environments by naturally looking around. In most VR systems, viewport orientation is directly controlled through one-to-one mapping between head pose and the virtual camera. While this mapping preserves a strong sense of spatial correspondence between physical and virtual motion, it shifts the burden of comfortable interaction onto the users. In conventional VR, users are often expected to be able to freely rotate their head and torso to explore the surrounding environment, or rely on external accommodations, such as swivel chairs or unrestricted space, to remain comfortable. This assumption breaks down in many real-world situations where large head movements are impractical, such as when users are seated, lying down, traveling, or in public settings where extensive movement may be constrained or socially awkward. As a result, not only is VR often fatiguing and uncomfortable, but its broader adoption is inherently constrained by where, for how long, and by whom it can be comfortably used.

To reduce the need for external accommodations and large physical movements, researchers have explored techniques that decouple viewport rotation from head motion. Controller-based techniques allow users to rotate the view using handheld devices, but they require manual input and occupy the hands, making them less suitable for hands-free interaction [49]. Head amplification techniques instead increase the virtual rotation produced by head movement, reducing the amount of physical motion required. However, amplification introduces a trade-off between comfort and controllability: Low gain provides only a limited reduction in physical effort. High gains reduce physical effort but increase sensitivity, causing large virtual rotations even with small or unintended head movements such as overshooting, disorientation, and simulator sickness [61]. This can make the viewing experience unstable and reduce the sense of spatial control.

In this work, we explore an alternative perspective motivated by the idea of Looking Around by Looking Around (LALA), which extends the viewport through natural eye movements. During visual exploration, the eyes typically move first to inspect peripheral regions before the head follows to bring those regions into central view. This eye-head coordination suggests that gaze can serve as a natural signal for extending the viewport. However, because gaze is primarily used for perception rather than control, designing a purely gaze-driven viewport control technique requires balancing responsiveness with perceptual stability.

Hence, the design of LALA was carefully guided by three principles aimed at ensuring user experience and comfort: (1) preserving stable viewing during natural gaze behavior, (2) enabling smooth and continuous viewport control for exploration, and (3) supporting rapid reorientation during large gaze shifts. To achieve this, LALA converts gaze signals into a control profile derived from the physiological and behavioral properties of human eye movements, and integrates this profile into a viewport control mechanism that accounts for both *oculomotor* and *perceptual* asymmetries.

We first construct an asymmetric omnidirectional gaze control profile that maps eye-in-head positions to control activation values (sub-section 3.1). We then translate these activation signals into viewport control formulations for yaw (subsubsection 3.2.1) and pitch (subsubsection 3.2.2) to account for perceptual asymmetries. Finally, as rotations are now decoupled from the HMD, we remap rotations and head movements to maintain natural viewing (subsection 3.3).

Together, these mechanisms of LALA produce an asymmetric omnidirectional control profile that enables natural, hands-free exploration while maintaining stable and comfortable viewing, allowing the eyes to steer the viewport toward any direction without constraining head movement. To our knowledge, LALA is the first to enable intuitive gaze-only pitch-yaw viewport control, achieved by accounting for oculomotor and perceptual asymmetries.

We evaluate LALA through controlled user studies comparing it with head amplification and baseline 1:1 mapping. Our results show that LALA achieved performance competitive to existing techniques while enabling fully hands-free interaction with minimal physical movement. Participants also strongly preferred LALA over our baselines, reporting favorable subjective experiences when using LALA to explore the virtual environment.

This work contributes:

- A gaze-based viewport control technique that enables stable omnidirectional VR exploration through natural eye movements alone, without requiring or obstructing movement of the head or torso.
- A control profile design for the eyes that accounts for both oculomotor and perceptual asymmetries and characteristics.
- An empirical evaluation demonstrating that LALA can support

• Hock Siang Lee is with Lancaster University. E-mail: h.s.lee3@lancaster.ac.uk

• Jinghui Hu is with Lancaster University. E-mail: j.hu23@lancaster.ac.uk

• Florian Weidner is with University of Glasgow. E-mail: florian.weidner@glasgow.ac.uk

• Haopeng Wang is with Lancaster University. E-mail: h.wang73@lancaster.ac.uk

• Hans Gellersen is with Lancaster University and Aarhus University. E-mail: h.gellersen@lancaster.ac.uk

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

practical exploration of VR environments with strong user preference and qualitative results better or competitive to existing head-based techniques.

2 VIEWPORT CONTROL TECHNIQUES

VR is often experienced without visibility of the physical environment and within limited real-world space, making intuitive and space-efficient travel techniques essential. These techniques include viewport control (controlling the virtual FOV) and locomotion (translating in virtual space), though research has focused more heavily on locomotion [10] as it is commonly assumed that users have enough space to control their view through rotation of head and torso. However, the range of head movement can be limited by posture [33, 48, 55], situation [4, 11] or ability [16, 38].

Alternative viewport control techniques are often designed to reduce user effort while improving navigation and accessibility [37], with a focus on modifying the horizontal yaw rotation. Within academia, head amplification techniques are a popular hands-free alternative, involving amplifying the movements of the head for users to reach further with head movement alone than physically possible [18, 43, 63]. Amplifying by a constant gain factor retains an absolute mapping between head pose and the viewport, supporting a user's sense of spatial orientation [41], but dynamic gain has also been explored [23, 27, 66]. A concern with head amplification is VR sickness due to the difference in visual and vestibular movement [39].

VR viewport control has often been explored using external controllers and sensors too, utilizing various body parts as a means of control. The most common techniques involves buttons, joysticks, or similar inputs on hand-held controllers [44]. Hands-free techniques that use body-leaning [30, 45, 58], shoulder-rotation [15], arm-movement [32] and more can also be found, but these approaches are typically designed for locomotion tasks.

There are also techniques that expand the user's effective FOV without interaction by using visual strategies such as scaling panoramic views to fit within a narrow FOV [2], layering multiple perspectives [46], or introducing pictures-in-picture [28].

Most recently, with eye-tracking now increasingly available in modern VR HMDs, academia have also recently started exploring using gaze for viewport control. Most notable for our work is how Lee et al.'s Gaze Gain and Gaze Pursuit techniques [25] harnesses different eye movements, even though they are 1D yaw-only techniques prioritizing performance (as opposed to comfort) and cannot be used for combined pitch-and-yaw rotations. Gaze Pursuit leverages smooth pursuit eye movements through a relative mapping, continuously rotating the viewport in the direction of the eye, with the rotation speed proportional to the eye-in-head angle. In contrast, Gaze Gain leverages saccadic eye movements with an absolute mapping, rapidly rotating the viewport toward an amplified combination of head and eye yaw angle, with rotation speed governed by a transfer function [25]. Both of these techniques utilized learnings from general gaze-based navigation, including scrolling and panning on desktops [47] and public displays [67, 68], and for 2D camera control in 3D video games [3, 56]. Though, works that have used gaze to manipulate the view in VR also exist, for example enabling gaze to orbit around an object [40], to subtly move the viewport during saccades for redirected walking [53], or to use peripheral gaze as a trigger for torso rotation [26]. Importantly, we highlight that extending gaze-based viewport control from yaw to omnidirectional pitch-yaw is not a straightforward dimensional extension, requiring a redesign of the control model itself (see subsection 3.1). Simple extensions would be difficult to use due to fundamental oculomotor (see subsection 3.1) and perceptual asymmetries (see subsection 3.2.2) between horizontal and vertical gaze.

3 LALA DESIGN

LALA is an omnidirectional gaze-based viewport control technique that allows users to stably explore VR environments by inducing additional viewport rotation toward the direction of gaze. Although gaze naturally reflects visual attention and user intent, the human visual system is optimized for perception rather than control. Consequently, effective

gaze-based viewport control must account for the characteristics and limitations of eye movements and visual perception.

An important consideration in omnidirectional viewport control is the oculomotor and perceptual asymmetries between the horizontal and vertical axes. *Oculomotor asymmetries* refer to the direction-dependent characteristics of eye movements, such as different directions requiring differing amount of effort. *Perceptual asymmetries* refer to our familiarity perceiving relative to a gravity-aligned reference frame and that vertical movements interact more strongly with gravitational and vestibular cues.

LALA accommodates for oculomotor asymmetries a asymmetric omnidirectional gaze control profile that maps eye-in-head positions to control activation values (subsection 3.1). We then translate these activation signals into distinct viewport control formulations for yaw (subsection 3.2.1) and pitch (subsection 3.2.2) to account for oculomotor asymmetries. Finally, as rotations are now decoupled from the HMD, we remap rotations and head movements to maintain natural viewing (subsection 3.3).

Notably, complexities of the human visual and oculomotor systems heavily influence LALA's design. While these complexities are well studied in vision science, they are less familiar in HCI and VR interaction research, so we detailed them together with the designs of LALA (as opposed to a related work section). We highlight that explicit use of vision science as a basis for interaction design is rare in HCI, and can be considered part of the contribution of LALA.

3.1 Asymmetric Omnidirectional Gaze Control Profile

LALA converts gaze data into an asymmetric omnidirectional control profile allowing easy interpretations of gaze signals, serving as the foundation for downstream viewport control. The goal of this profile is to capture the key characteristics of human eye movements so that gaze signals can be interpreted in a way that supports stable viewing, smooth continuous control and rapid reorientation. To achieve this, the control profile is derived from the behavioral and physiological properties of fixations, smooth pursuit, and saccadic eye movements.

To design such a profile, we first consider how gaze has been used in general interactions. Large eye-in-head angles have been used as an indicator of the user's intent to interact with interface elements, as they often reflect effortful gaze shifts since natural gaze rarely enters such peripheral regions [7, 48]. These interactions often use a trigger-based binary decision. However, smooth viewport control requires a continuous control signal to avoid jerky start-stop behavior caused by eye movements such as the optokinetic reflex. This control signal is most naturally defined as an absolute mapping of eye-in-head angles, rather than relative velocity or gestures. Eye positions inherently encode spatial offsets, which users can interpret through proprioception and their understanding of the surrounding environment, making the relationship between eye angle and control signal more intuitive and easier to learn compared to relative mappings.

Having established that viewport control should be driven by a continuous mapping of eye-in-head angles, the next question is how various eye movements affect this mapping. Fixations, which typically occur near the central eye-in-head region, involve aligning a stationary object or area of interest with the fovea for at least ~ 200 ms to enable detailed visual processing [22]. Hence, to maintain stable viewing within the central eye-in-head region, the viewport should remain largely stationary, corresponding to a control signal close to zero.

Outside of this central region, while gaze remains comfortably reachable by the eyes alone, scene rotations induced by viewport control would cause fixations to naturally transition into smooth pursuits. Smooth pursuits are particularly important for viewport control, as they enable users to continuously track objects and maintain engagement with displayed content while the viewport is moving. However, smooth pursuit can reliably track motion only up to approximately $\sim 30^\circ/\text{s}$ before catch-up saccades are triggered [36]. While saccades are much faster, their ballistic nature makes their endpoints almost impossible to adjust once initiated [21]. Consequently, the ideal control signal for viewport rotational speeds must be constrained while within this region. If the viewport moves too quickly, smooth pursuit will frequently trig-

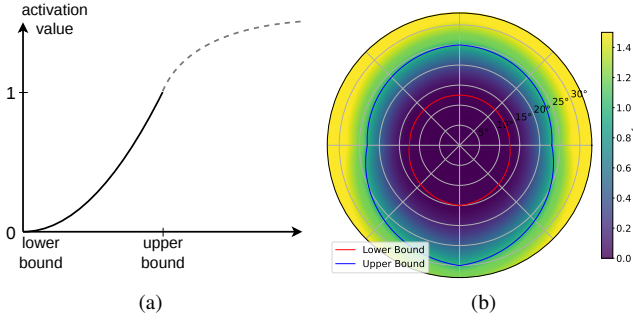


Figure 1: (a) Activation curve between a lower and upper bound from a slice of (b) which is the full 2D mapping between eye-in-head directions and activation values that accounts for anatomical asymmetries of the eye across axes.

ger catch-up saccades, and saccades may miss their intended landing positions as visual targets shift during their execution. This results in viewport motion that appears stuttery and difficult to control [25], potentially increasing cybersickness and disorientation [50, 51].

Further out represents the peripheral eye-in-head region, which the eyes can reach but not comfortably linger in. This region is rarely accessed during natural gaze and is typically only reached during large-amplitude gaze shifts that are assisted by head movements [48]. Such gaze shifts can even produce eye-in-head angles that exceed modern VR HMDs’ FOVs, for example, when tracking a fast moving object and our visual system predicts a future position outside the FOV. Because these saccades are often guided top-down rather than toward a currently visible stimulus, the constraint limiting viewport speed to prevent saccades from missing its intended target is less critical. Instead, the limiting factor for viewport speed is our ability to process newly appearing scene content and deciding when to stop or adjust gaze. Consequently, the control signal in the peripheral region should increase substantially to enable rapid reorientation, but remain bounded to avoid overwhelming visual processing, which would degrading user control [25].

Overall, the control signal can be conceptualized as having three regions where interactions are primarily from fixations, smooth pursuits, or saccades. It stays near zero in the central region to support stable fixations, rises gradually through the mid-range to accommodate smooth pursuits, and increases sharply in the peripheral region, where it is ultimately bounded by visual processing limits. To capture this behavior, we use a sigmoidal formulation, which ensures smooth transitions between regions while constraining the control signal at extreme eye-in-head angles.

For the early and mid-range response of the sigmoid — representing the central and still-comfortable regions of the eye — we hypothesize that the muscular effort required to maintain peripheral eye positions correlates with the strength of the user’s intent to interact with the viewport. Accordingly, we base this portion of the curve on measurements of eye-muscle tension required to maintain *horizontal* eye positions [9]. To ensure smooth behavior beyond this range (>1) — representing the peripheral, sometimes still-reachable regions of the eye — we extend the curve using Hermite interpolation as it accounts for the curve derivative, allowing the curve to more smoothly flatten. We intentionally used values >1 so that the 0–1 range could remain well supported by vision science, whereas eye behavior at extreme angles is less understood. Extending beyond this range also accommodates measurement noise, inter-individual variability, and eye-in-head angles that may exceed the HMD’s FOV. Combined, this produces an overall sigmoid-like curve as seen in 1a. The lower asymptote at $y = 0.01$ was motivated by pilot testing, which suggested that hard stops were less intuitive for some users, while the upper asymptote at $y = 1.5$ was selected as a pragmatic extension.

However, oculomotor asymmetry prevents simple extrapolation of the control signal from horizontal-only to omnidirectional. Our visual

and oculomotor system does not treat the vertical pitch axis symmetrically with the horizontal yaw axis [5, 13, 20]. For instance, smooth pursuits are inherently asymmetric, with performance generally superior for horizontal yaw compared to vertical pitch [13]. These oculomotor asymmetry affect eye-muscle tension, which in turn affects preferred eye-in-head amplitudes across gaze bearings.

To account for this, the control signal was scaled to fit between the lower and upper bounds of preferred eye-in-head amplitudes observed during gaze shifts towards different bearings [48]. We assume the overall shape of the sigmoid (i.e., the general relationship between eye-in-head angle and effort) is similar across bearings, and that direction-dependent differences are sufficiently accounted for by the directional bounds. This results in the 2D pitch-and-yaw representation depicted in 1b. Together, these bounds create a 2D control profile that may be perceived as comprising three regions: a central area for stable viewing, an intermediate region supporting comfortable passive viewport control, and a peripheral zone enabling active viewport control.

The resulting activations values from the control profile is defined with respect to eye-in-head space, which can be decomposed into components along the relevant world axes using basic trigonometry. This is particularly important as eye rotations do not align with world axes when head roll is present. All in all, our control profile accounts for many of the complexities of the eye, which has the additional benefit of decoupling and simplifying downstream technique design from such complexities. The representative values [9, 48] are provided together with the source code in the Appendix.

3.2 Perceptual Asymmetry: Differences in Pitch and Yaw Control

The gaze control profile described above produces activation values in eye-in-head space, which must be translated into viewport rotations to enable gaze-driven navigation. However, this translation cannot be applied uniformly across directions. Human perception treats horizontal (yaw) and vertical (pitch) rotations differently, largely due to gravity and vestibular cues. As a result, LALA adopts different control formulations for yaw and pitch: yaw is implemented using a relative control, while pitch is implemented using an absolute mapping aligned with gravity.

3.2.1 Relative Yaw Control via Gaze

For yaw, our goal is to enable easy 360° surround viewing. Among the yaw-only gaze-based techniques reviewed in section 2, the Gaze Pursuit technique [25] achieved the strongest performance, so we draw inspiration from it.

In both LALA and Gaze Pursuit, the yaw of the viewport continuously rotates in the direction of the eye in the horizontal direction. As an object of interest comes into view, the eye naturally fixates and smooth pursuits back towards the centre.

$$\text{velocity}_{\text{yaw}} = m \times \text{activationValue}_{\text{horizontal}} \times (\text{eyeInVRWorld}_{\text{yaw}} - \text{currentDirection}_{\text{yaw}}) \quad (1)$$

However, both techniques differ in how they determine the viewport angular velocity. LALA’s yaw velocity (see Equation 1) is proportional to the activation value (see subsection 3.1) and $\text{eyeInVRWorld}_{\text{yaw}}$ (the angle between the current viewport direction and the current gaze direction in VR). m is a tunable parameter controlling interaction sensitivity, analogous to cursor sensitivity in pointing interfaces, and was set to 4 based on pilot testing to achieve stable and usable behavior.

3.2.2 Absolute Pitch Control via Gaze

For pitch, our goals are to enable easy and intuitive $\pm 90^\circ$ pitch control. However, unlike yaw, this is more complicated due to perceptual asymmetries.

During natural horizontal gaze shifts (yaw), gravity remains aligned with the head’s orientation, meaning gravity-sensitive receptors (e.g., the otoliths) are minimally affected. In contrast, pitch is more strongly affected by the continuous feedback from these gravity-sensitive receptors [20, 59], as opposed to the current view direction. This causes

perceptual asymmetry between yaw and pitch, whereby yaw is perceived in a relative frame of reference, while pitch is perceived in an absolute frame of reference relative to gravity.

Hence, a naive relative mapping for pitch control is unintuitive. It makes all vertical motion equally effortless, allowing users to rotate past natural head movement limits — looking fully up and down — without resistance, easily rotating into being “upside-down”. Yet, users are unlikely to want or benefit from such orientations, as humans have limited spatial familiarity and reduced sensory resolution when presented with upside-down sceneries [8, 20]. Therefore, viewport pitch control is ideally implemented using an absolute mapping, with reference to human-perceivable cues such as gravitational or proprioceptive signals.

$$\text{goalDirection}_{\text{pitch}} = g \times \text{eyeInWorld}_{\text{pitch}} \quad (2)$$

In LALA, we employ an absolute mapping between viewport pitch and eye-in-world angle (Equation 2) with an amplification factor $g = 2$ to obtain a “goalDirection”, representing where the viewport should pitch to. An amplification of 2 allows the technique to cover the entire $[-90^\circ, 90^\circ]$ pitch range using only eye movements. Combined with the relative yaw mapping, it means upright viewing towards any direction is possible using only eye movements.

$$\begin{aligned} \text{velocity}_{\text{pitch}} = & m \times \text{activationValue}_{\text{vertical}} \\ & \times (\text{goalDirection}_{\text{pitch}} - \text{currentDirection}_{\text{pitch}}) \end{aligned} \quad (3)$$

The viewport pitch velocity towards this “goalDirection” is governed by Equation 3, where m is a tuning constant controlling interaction sensitivity, set to 1.8 based on pilot testing to achieve stable and usable behavior. This means LALA’s pitch can be interpreted as a “delayed” absolute mapping, where its pitch velocity is proportional to both the activation value (see subsection 3.1) and the angular difference between the current viewport pitch and the amplified gaze pitch. A “delayed” approach is necessary here to prevent disorientation from saccades missing their intended landing positions if visual targets shift too much during their execution.

This design improves intuitiveness (by aligning viewport movement to gravity, allowing a “sense” of up/down during movement) while preventing users from unintentionally disorienting themselves or flipping the scene upside down (by using an absolute mapping for pitch tuned towards covering the $\pm 90^\circ$ pitch range). We also added a $1^\circ/\text{s}$ rotation towards 0 pitch if the viewport pitch is within $\pm 10^\circ$ to help realignment back to the horizon and sense of orientation.

Notably, the yaw and pitch velocity equations in LALA share the same structure, differing only in the reference direction used to compute the angular difference relative to the current viewport direction. This highlights the advantage of subsection 3.1, which abstracts eye movement complexities from technique design. This allows the remainder of LALA to remain agnostic to the underlying behaviour of the eyes, making a complete technique redesign between axis unnecessary. Furthermore, both velocity formulas intentionally use homogeneous equations containing no constant terms, ensuring that small to moderate eye-in-head angles do not induce viewport rotations. In other words, while the eye remains within a comfortable range, the viewport behaves perceptibly the same as the default 1:1 HMD mapping, ensuring that LALA remains effectively invisible until an “uncomfortable” gaze shift occurs.

3.3 Remapping Rotations and Head Movement

In LALA, viewport orientation is partially decoupled from the HMD because gaze can rotate the view independently of head motion. This decoupling introduces ambiguities in how head rotations should affect the final viewport rotation. Under standard quaternion composition, yaw and pitch interact depending on the current orientation of the viewport. As a result, head rotations could produce different perceived motions depending on where the user is looking, leading to inconsistent yaw and pitch behaviour. To avoid this issue, we remap head and viewport rotations to preserve perceptually consistent control.

To address this, LALA (see Equation 3 and Equation 1) uses an Euler-angle-like mapping which preserves the perceptual separation between yaw and pitch, which would otherwise break under quaternion composition. The mapping intentionally re-introduced gimbal lock when looking directly up or down to reflect natural limits of human head rotation and visual search behavior under gravity. All head and eye yaw rotations are remapped to only alter the viewport (global) yaw regardless of current viewport pitch; and all head and eye pitch rotations are remapped to only alter the viewport (global) pitch regardless of the current viewport yaw. This behavior is consistent with known constraints on biological eye-head coordination such as Listing’s law [17, 65].

This remapping better matches natural head movement and how rotations are interpreted in egocentric reference frames. For example, when looking upward while spinning in an office chair, perceived pitch (looking up/down) and yaw (turning left/right) remain consistent relative to the body and gravity-defined vertical. In contrast, aircraft-style rotation models tilt the entire reference frame, causing subsequent yaw rotations to behave differently. By aligning head and viewport movements with the body and gravity-aligned vertical, the interaction feels more intuitive [1].

Likewise, head displacements are remapped to the current viewport orientation, so moving the HMD in the direction it faces moves the viewport in the direction it faces, regardless of their relative orientations. This allows users to, for example, gain elevation in VR by pitching the viewport upward and walking forward. Although our study did not elicit large head displacements, this remapping remained necessary because natural head and torso rotations produce small but non-negligible HMD and eye displacements.

In summary, LALA combines four modular components — asymmetric omnidirectional gaze control profile, a yaw control, a pitch control, and head-movement remapping — to enable natural, omnidirectional viewport control with minimal physical effort. Together, these modules allow the technique to remain mostly unobtrusive during comfortable eye movements, while supporting intuitive navigation when larger gaze shifts occur.

4 EVALUATION

To evaluate our viewport control techniques across diverse conditions, we conducted a user study consisting of two tasks and compared them to existing baselines. We selected these tasks as they represent contrasting user behaviors with distinct gaze patterns and cognitive demands: efficient target acquisition where the direction is known and predictable, versus open-ended visual search for targets at unknown and unpredictable directions.

The first is a modified Fitts’ law task (see subsection 4.1) where participants looked for a target at a predictable and known direction, allowing them to move efficiently towards the current target while planning their movements accordingly for future targets. The second is a visual search task (see subsection 4.2) where participants freely searched their surroundings for a target at an unknown direction, necessitating them to scan their environment without prior knowledge of target locations.

In both tasks, targets are selected via eye-gaze combined with a manual button press, allowing participants to have a fast, immediate, explicit way to confirm when the target is perceived to have aligned with their gaze. The design choice of using button selection as opposed to gaze-dwell selection, which is common in gaze-based interactions, is due to our focus on evaluating viewport alignment performance rather than selection mechanisms. By requiring an intentional button press, we avoid the Midas Touch problem in which unintended gaze fixations trigger false selections, while also eliminating effects caused by different gaze dwell durations.

The targets appear 10m away from the participant and are 15 visual degrees in size (the approximate size of macular vision), making it easy to select even in the presence of eye-tracking inaccuracies. Combined with the button-press selection method, this design emphasizes viewport alignment — bringing targets into view — rather than precise selection accuracy.

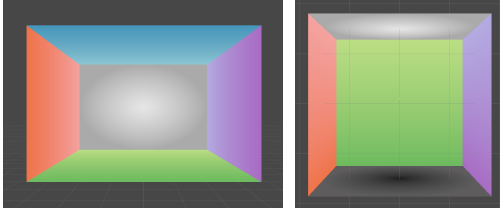


Figure 2: Cube environment for both our tasks viewed from the front (left image) and top (right image). Image proportions modified for better sense of perspective and orientation.

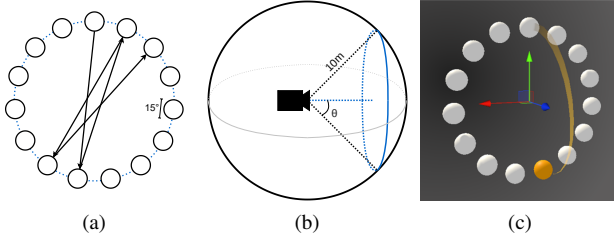


Figure 3: Our Alignment Task targets positions viewed from the front as a layout schematic (a), and the side (b). 15 targets of size 15° visual angle were placed along the blue ring 10m away, where θ was the radius $\{45^\circ, 90^\circ, 135^\circ\}$. (c) shows the study setup when θ is 90° , positioning all the sphere targets around the VR camera (represented as the multicolored axis arrowhead) such that they all lie on the same plane. The orange sphere represents the target to be selected, the white spheres are only shown for figure purposes. The orange arc provides visual feedback connecting the previous target to the next.

To give users a clear sense of orientation during both tasks – and to avoid the perception that targets were moving rather than the viewport – it was necessary to include sufficient visual cues in the environment. However, visual cues and distractions could introduce noise into the data. To balance these needs, both tasks were conducted in a large virtual cube ($30 \times 30 \times 30$ m) with walls in distinct colors, each featuring a subtle gradient that becomes lighter in the forward-facing direction, as shown in Figure 2.

4.1 Alignment Task: Looking towards known directions

Fitts’ law experiments are widely used to test and model interactions that involve efficiently pointing rapidly towards a known and predictable target before selecting it. Our study draws inspiration from Fitts-style target acquisition tasks to evaluate how effectively users can control the viewport to locate targets in known directions, both within and beyond the initial field of view. Importantly, the task focuses on bringing targets — often initially from outside the FOV — into central vision rather than rapid target acquisition, distinguishing the task from traditional Fitts’ law paradigms.

In our task, the targets were positioned as shown in Figure 3 and selected via gaze and a button press. We tested ring radiuses (θ) of 45° , 90° , and 135° , selected to represent a range of head-eye coordination demands. 45° was initially within FOV but not comfortably reachable with the eyes alone; 90° was initially outside FOV but comfortably reachable via head movements; and 135° was both initially outside FOV and also not comfortably reachable via head movements. Each ring contains a relatively high number of targets (15), providing a higher angular sampling density to capture the fine-grained, directional characteristics of the interaction techniques under investigation.

In the design of traditional Fitts’ Law pointing tasks, to ensure the movement is efficient and direct, targets are always kept in view to provide essential visual feedback for corrective (closed-loop) sub-movements to occur during selection [34]. However, the targets in our task have large angular separation (θ) and are mostly not in the FOV.

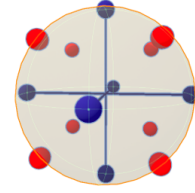


Figure 4: Our Search Task targets positions. 14 targets of size 15° visual angle (2.63 m wide) were placed 10 m away. Target were placed in directions that can be thought of as extending from the center of a cube to the center of each face (blue) and to each corner (red). Colors and lines are for purely illustrative reasons.

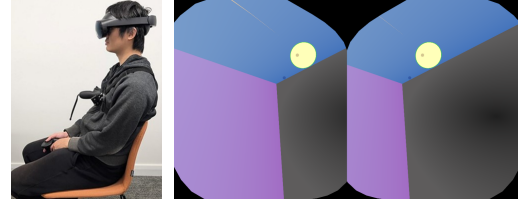


Figure 5: A study participant wearing the chest harness while seated on a stationary chair for the study (left), and what they were seeing in VR (right). The images was taken during the Alignment task when the radius θ was 135° .

To maintain visual feedback and help participants locate the targets – and to discourage “shortcuts” (e.g., rotating in the shorter direction when $\theta > 90^\circ$), a visual guide is displayed connecting the previous and next targets. The visual guide, shown in 3c as the orange arc, can be interpreted as the straight line path traveled between targets in a standard Fitts’ Law task, projected onto a sphere centered on the participant to form an arc. This line is not the shortest arc between targets (i.e., the geodesic), but is part of the circular intersection between a flat plane and the sphere, where the plane intersects the center of both targets, such that the arc spans exactly θ degrees.

4.2 Search Task: Looking towards unknown directions

In our alignment task, the knowledge of target positions guides gaze and viewport control behavior. However, in real-world scenarios, the locations of points of interest are not necessarily known beforehand, requiring a search phase before interaction. We therefore adapted a self-directed visual search task to evaluate how effectively users can control the viewport to search and locate targets in unknown directions beyond the initial FOV. In our task, targets appear one at a time at fixed positions (shown in Figure 4), with each new target displayed only after the previous one was selected. Positions are selected randomly with two constraints: no position can appear consecutively, and each position appears exactly three times total.

While self-directed visual search beyond the initial FOV in VR has been studied primarily in the context of optimising search strategies and reducing search time [14, 31, 52], our objective differs. Rather than focusing on search strategy and efficiency, we wanted to investigate whether exploring all directions can be achieved with minimal effort. To this end, the target positions (shown in Figure 4) were deliberately arranged so participants were required to search in all directions (including directly behind, above and below), thereby ensuring completeness of exploration.

4.3 Procedure

At the start of the task, the experimenter welcomed the participant and explained the purpose and procedure of the task. After providing consent, participants filled out a demographics questionnaire asking for age, gender, vision (normal, corrected with glasses, corrected with contact lenses), experience with video games (never, rarely, monthly, weekly,

daily), experience with VR (never, rarely, monthly, weekly, daily), and experience with eye tracking (never, rarely, monthly, weekly, daily). Then, participants were instructed to put on a provided chest harness with sensors attached for torso tracking (shown in Figure 5). Participants were seated in a stationary, non-rotating chair, and eye-tracking calibration was performed at each headset donning. Participants completed both tasks with one technique at a time before moving to the next. Technique order (balanced Latin square) and task order within each technique were counterbalanced across participants.

Before using each technique, participants received a brief explanation of the technique from the experimenter. Before the start of each task, participants practiced using the technique via a practice task – a greatly shortened version of the task consisting of 15 selections. Participants were prompted at the end of each practice task if they wanted to repeat the practice or to continue with the study task. Most participants chose not to continue practicing, with a rare few choosing to practice an additional time. No participants chose to continue practicing more than once. After completing each task, participants took off the HMD and were given as much time as necessary to rest and complete questionnaires regarding the techniques. The task lasted under 70 minutes.

4.4 Apparatus

We implemented and ran the task with Unity 2021.3.27f1 on a computer with an Intel Core i7-12700 CPU, 16 GB RAM, and an NVIDIA GeForce RTX 3070 Ti GPU, connected to a Meta Quest Pro VR HMD with a Quest Link cable. The Meta Quest Pro had a 111.24° diagonal FOV, 1800x1920 pixels resolution per eye, and 90 Hz refresh rate. Torso tracking used a Meta Quest Pro handheld controller attached to the front of a chest harness.

4.5 Techniques

The techniques compared in our tasks were:

- 1xHead (baseline) represents the default 1:1 mapping commonly found in VR experiences, where HMD rotation directly maps to viewport rotation.
- 2xHead is a HeadGain implementation using our head remapping module (see subsection 3.3) that amplifies head rotation by a factor of 2, doubling the HMD yaw and pitch such that the viewport has twice the amount of yaw and pitch (relative to the forward direction).
- LALA is our gaze-based viewport control technique introduced in section 3.

Initially, we considered comparing LALA against straightforward pitch-and-yaw extensions of Gaze Gain and Gaze Pursuit. However, oculomotor and perceptual asymmetries (see subsection 3.1 and subsection 3.2.2) prevented these originally 1D yaw-only techniques from being readily extended to combined pitch–yaw rotations. Moreover, naive extensions yielded unfair comparisons: pilot studies showed that pitch-and-yaw Gaze Pursuit produced disorienting viewing, while pitch-and-yaw Gaze Gain caused excessive eye strain. We also considered unmodified yaw-only Gaze Gain and Gaze Pursuit, but these were unsuitable because our targets vary substantially in pitch, making pitch equivalent to the 1xHead baseline, while yaw performance had already been evaluated in [25]. Instead, we included 2xHead as a reference condition because HeadGain techniques are common in prior work, and higher gains such as 2x have not been evaluated for our target scenario.

4.6 Measures

Our studies aimed to assess how comfortably and effectively participants could control the viewport to view targets. For quantitative measures, we recorded the cumulative angular rotation and range of motion of the eyes, head, and torso for each selection. We also measured the time taken for the participant’s gaze to first land on the target (i.e., completion time) and the error rate (i.e., missed selections).

For subjective measures, participants rated the following statements on a 7-point Likert scale: “The technique was easy to use,” “The technique was comfortable to use,” and “You had a good sense of control

over the technique”. These ratings provided an overall indication of ease of use, comfort, and perceived controllability. Participants also completed the RAW NASA-TLX [6] and VRSQ [19] questionnaires to capture perceived workload and cybersickness, respectively. Finally, participants were invited to rank the techniques according to their preference and share any additional comments.

4.7 Sample

We initially recruited 21 participants from the local university. However, three were excluded from the final analysis: one participant dropped out immediately due to motion sickness unrelated to the task or technique, another experienced issues with eye-tracking calibration, and the third had an eye injury affecting upward gaze in the right eye. This resulted in a final sample of 18 participants (8 identified as male, 10 as female). The average age was 29.0 years (SD = 8.73; range = 22–53). Vision was reported as normal (n = 5), corrected with contact lenses (n = 2), or corrected with glasses (n = 11). Video gaming frequency was reported as daily (n = 4), weekly (n = 3), monthly (n = 2), less than once a month (n = 7), or never (n = 2). Both VR usage and eye-tracking usage were reported as none (n = 3), rarely (n = 12), or monthly (n = 3). All participants received a £10 Amazon voucher. The university’s Institutional Review Board approved the study.

5 RESULTS

Quantitative objective data of both tasks were analyzed using two-way repeated measures ANOVA ($\alpha = .05$). When the assumption of sphericity was violated, as indicated by Mauchly’s test, Greenhouse-Geisser corrections were applied (which can cause non-integer degrees of freedom). Normality was assessed using the Shapiro-Wilk test and Q-Q plots. If the assumption of normality was violated, the Aligned Rank Transform (ART) procedure [64] was used. Post-hoc pairwise comparisons were performed only in cases of statistically significant main or interaction effects and were limited to within-factor comparisons (i.e., conditions differing along a single factor only). Bonferroni correction was applied where appropriate. Pairwise *t*-tests were used when parametric assumptions were met, whereas ART-C were applied when assumptions were violated. Effect sizes are reported as general eta squared (η_g^2) when parametric assumptions were met, partial eta squared (η_p^2) when ART was applied. Subjective data were analyzed using Friedman tests, with effect sizes reported as Kendall’s W. In all bar charts, only comparisons that are not statistically significant are labeled (ns). Statistical significance was prioritized in the main text. Further statistical details and tables summarizing the statistics for easier viewing are provided in the Appendix.

5.1 Alignment Task Results

Table 1 and Table 2 summarizes the Main and Interaction Effects statistics for our Alignment Task.

5.1.1 Completion Time

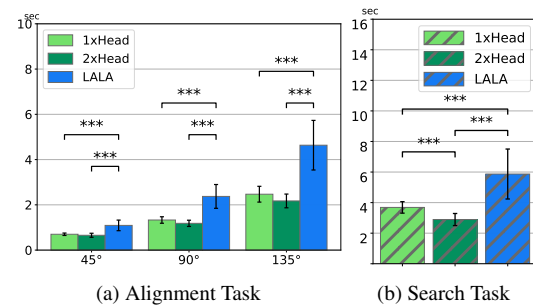


Figure 6: Completion Time Per Trial, measured as time taken to gaze at the target, for our Alignment Task by Radius and Technique (a), and for our Search Task (b). Error bars represent the 95% confidence interval. Lower values are better.

Technique × Radius Interaction Effect				Technique Main Effect				Radius Main Effect			
	<i>F</i> value	<i>p</i>	η^2	<i>F</i> value	<i>p</i>	η^2		<i>F</i> value	<i>p</i>	η^2	
Alignment Speed	F(4,136) = 19.532	<.001	.364 ***	F(2,136) = 93.232	<.001	.578 ***		F(2,136) = 332.998	<.001	.830 ***	
Error Rate	F(4,136) = .485	.747	.014	F(2,136) = .721	.488	.001		F(2,136) = .773	.463	.011	
Cumulative Rotation of	Eye	F(1.85,31.47) = 16.153	<.001	.155 ***	F(2,34) = 36.385	<.001	.429 ***	F(1.22,20.83) = 256.171	<.001	.63 ***	
	Head	F(1.73,29.34) = 31.407	<.001	.241 ***	F(1.13,19.21) = 63.309	<.001	.601 ***	F(1.33,22.61) = 346.326	<.001	.64 ***	
	Torso	F(4,136) = 3.086	<.018	.083 **	F(2,19.21) = 63.309	<.001	.601 ***	F(2,136) = 36.846	<.001	.351 ***	
Max Pitch of	Eye	F(4,68) = 23.082	<.001	.117 ***	F(2,34) = 36.385	<.001	.429 ***	F(2,34) = 57.033	<.001	.187 ***	
	Head	F(1.57,26.69) = 8.995	<.001	.084 ***	F(1.36,23.15) = 83.400	<.001	.627 ***	F(1.47,24.92) = 77.505	<.001	.258 ***	
	Torso	F(4,136) = 2.605	<.038	.071 *	F(2,136) = 3.015	.052	.050	F(2,136) = 3.527	<.003	.049 **	
Max Yaw of	Eye	F(2.69,45.66) = 4.804	<.001	.039 ***	F(2,34) = 34.151	<.001	.410 ***	F(2,34) = 125.892	<.001	.408 ***	
	Head	F(1.85,31.50) = 34.883	<.001	.028 ***	F(1.18,19.98) = 60.201	<.001	.606 ***	F(1.40,23.81) = 259.584	<.001	.614 ***	
	Torso	F(4,136) = 10.740	<.001	.240 ***	F(2,136) = 43.026	<.001	.388 ***	F(2,136) = 53.704	<.001	.441 ***	

Table 1: Alignment Task: ANOVA statistics. Effect sizes reflect η_g^2 when parametric assumptions were met, and η_p^2 when ART was applied.

		Technique Main Effect			
		Friedman Stat	<i>p</i>	<i>W</i>	
NASA	Physical Demand	$\chi^2(2) = 16.6$	<.001	.461	***
TLX	Effort	$\chi^2(2) = 7.44$	.024	.207	*

Table 2: Alignment Task: Significant Friedman statistics for our subjective data. Effect sizes reported as Kendall's *W*. Full statistics can be found in the Appendix.

There was a significant Technique × Radius interaction effect, indicating that Technique affected Completion Time (i.e., time taken to gaze at the target) differently at different Radius levels. Post-hoc pairwise analysis (indicated in 6a) revealed significant differences between all Technique pairs within each Radius level (all $p \leq .001$) except for the (1xHead, 2xHead) pair (all $p \leq 1$). It also revealed significant differences between all Radius pairs within each Technique condition (all $p \leq .001$).

There were also significant main effects of both Technique and Radius — though qualified by the significant interaction effect. Post-hoc pairwise analysis revealed significant differences between all Techniques (all $p < .001$) and all Radius levels (all $p < .001$).

5.1.2 Error Rate

There were no significant Technique × Radius interaction effects nor any significant main effects of Technique or Radius, indicating that Error Rate (i.e., missed selections) was similar across all conditions. The overall mean Error Rate across all conditions was $\mu = 0.070$ (95% CI = ± 0.015).

5.1.3 Cumulative Eye, Head, and Torso Rotations

There was a significant Technique × Radius interaction effect for Cumulative Eye, Head, and Torso Rotations (see Table 1), indicating that the effect of Technique varied by Radius level. Post-hoc pairwise comparisons were conducted between Techniques within each Radius and between Radius levels within each Technique showed that: For Cumulative Eye Rotation, all Technique pairs differed significantly at each Radius (all $p < .019$), and all Radius levels differed within each Technique (all $p < .001$); For Cumulative Head Rotation, all Technique pairs differed significantly within each Radius (all $p < .001$) except 2xHead vs LALA ($p = 1$), and all Radius levels differed within each Technique (all $p < .007$); For Cumulative Torso Rotation, no significant differences were found between Techniques or Radius levels (all $p = 1$).

Significant main effects of Technique and Radius were also found for all three rotation measures. Post-hoc pairwise analysis revealed that: For Cumulative Eye Rotation, all Technique and Radius comparisons were significant (all $p < .001$); For Cumulative Head Rotation, all were significant (all $p < .001$) except 2xHead vs LALA ($p = .080$); For Cumulative Torso Rotation, all were significant (all $p < .001$) except 2xHead vs LALA ($p = 1$).

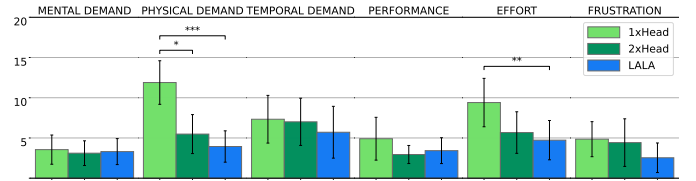


Figure 7: Alignment Task: Raw NASA TLX by Technique. Error bars represent the 95% confidence interval. Lower values are better.

5.1.4 Subjective Stats

Analysis of the subjective data collected in our Alignment Task revealed a significant effect of Technique in Physical Demand and Effort, with only non-significant small effects of Technique on other metrics (all $p > .104$). Post-hoc Wilcoxon signed-rank tests revealed that 1xHead had significantly higher Physical Demand compared to both 2xHead ($V=139$, $p=.010$) and LALA ($V=165$, $p<.001$). Additionally, 1xHead also had significantly higher Effort compared to LALA ($V=151$, $p=.008$). There was no significant effect of technique on all VRSQ subscales (all $p > .683$). Overall means ($\pm 95\%$ CI): Oculomotor Disturbance ($M = 17.13 \pm 4.82$), Disorientation ($M = 11.11 \pm 5.09$), and Total Simulator Sickness ($M = 14.12 \pm 4.53$).

For the remaining subjective metrics, there was no significant effect of Technique (all $p > .104$), with all effect sizes classified as small. Full subjective data can be found in the Appendix.

5.2 Search Task Results

		Technique Main Effect		
		<i>F</i> value	<i>p</i>	η^2
Alignment Speed		F(2,34) = 41.43	<.001	.709 ***
Cumulative Rotation of	Eye	F(1.28,21.73) = 26.422	<.001	.473 ***
	Head	F(2,34) = 36.58	<.001	.683 ***
Max Pitch of	Eye	F(2,34) = 28.942	<.001	.630 ***
	Head	F(1.2,20.39) = 43.012	<.001	.569 ***
Max Yaw of	Eye	F(2,34) = 23.569	<.001	.322 ***
	Head	F(2,34) = 43.932	<.001	.721 ***

Table 3: Search Task: Significant ANOVA statistics. Effect sizes reflect η_g^2 when parametric assumptions were met, and η_p^2 when ART was applied.

Table 3 and Table 4 summarizes the Main Effects statistics for our Search Task.

5.2.1 Completion Time

Analysis of Completion Time, measured as time taken to gaze at the targets, revealed a significant effect of Technique. As seen in 6b, post-hoc pairwise analysis revealed that 2xHead was significantly faster than 1xHead ($\hat{\Delta} = 12.389$, $SE = 2.740$, $t(34) = 4.521$, $p < .001$) which in turn was significantly faster than LALA ($\hat{\Delta} = 12.556$, $SE = 2.740$, $t(34) = 4.582$, $p < .001$). Furthermore, 2xHead outperformed LALA

		Technique Main Effect			
		Friedman Stat	p	W	
NASA	Physical Demand	$\chi^2(2) = 14.30$	<.001	.398	***
	Performance	$\chi^2(2) = 7.30$	.026	.203	*
	Effort	$\chi^2(2) = 8.48$	.014	.236	*
	Frustration	$\chi^2(2) = 7.97$	.019	.221	*

Table 4: Search Task: Significant Friedman statistics for subjective data. Effect sizes reported as Kendall's W. Full statistics can be found in the Appendix.

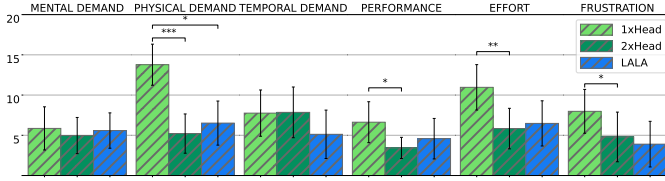


Figure 8: Search Task: Raw NASA TLX by Technique. Error bars represent the 95% confidence interval. Lower values are better.

($\hat{\Delta} = 24.944$, $SE = 2.740$, $t(34) = 9.102$, $p < .001$), having the fastest Completion Time overall.

5.2.2 Error Rate

Analysis of Error Rate revealed no significant effect of Technique. The overall mean Error Rate across all conditions was $\mu = 0.070$ (95% CI = ± 0.014).

5.2.3 Subjective Stats

Analysis of the subjective data collected in our Search Task revealed a significant effect of Technique in Physical Demand, Performance, Effort, and Frustration. Post-hoc Wilcoxon signed-rank tests revealed that 1xHead was significantly worse compared to 2xHead in terms of Physical Demand ($V=165$, $p<.001$), Performance ($V=118$, $p=.029$), Effort ($V=155$, $p=.004$), and Frustration ($V=118$, $p=.032$). Additionally, 1xHead also had significantly worse Physical Demand compared to LALA ($V=150$, $p=.002$). There was no significant effect of technique on all VRSQ subscales (all $p>.059$). Overall means($\pm$ CI): Oculomotor Disturbance ($M = 23.77 \pm 5.78$), Disorientation ($M = 16.42 \pm 5.78$), and Total Simulator Sickness ($M = 20.09 \pm 5.32$).

For the remaining subjective metrics, there was no significant effect of Technique (all $p>.126$), with all effect sizes classified as small. Full subjective data can be found in the Appendix.

5.3 User Preference

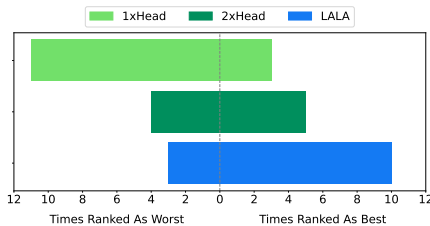


Figure 9: Times each technique was voted as Best or Worst

Figure 9 presents participants' ratings of the best and worst techniques after experiencing all options. LALA was rated as the best technique by the most participants (10), followed by 2xHead (5), and 1xHead (3). Conversely, 1xHead was rated as the worst technique by the most participants (11), followed by 2xHead (4), and LALA (3). Notably, LALA was rated the best technique by an absolute majority, whereas 1xHead was rated the worst by an absolute majority.

6 DISCUSSION

In this work, we compared three viewport control techniques — 1xHead, 2xHead, and LALA — to examine how different control strategies affect performance, effort, and user preferences. Across both tasks, LALA was slower but reduced viewing effort compared to the baseline 1xHead and was the most preferred technique overall. 2xHead was fast but was less preferred, with similar taskload results. When directly comparing LALA and 2xHead, participants preferred LALA despite its slower performance. Both techniques lead to similar levels of effort and error. These results highlight a trade-off between the natural 1:1 mapping (1xHead), performance-oriented head amplification (2xHead), and gaze-driven low-effort viewing (LALA).

6.1 LALA's tradeoff: preference vs speed

The results reveal a clear performance spectrum across the three techniques. The baseline 1xHead relies on direct head-to-viewport mapping, providing natural control but requiring large physical rotations that increase physical demand, forcing extreme head movements, ultimately leading to reduced preference. 2xHead amplifies head motion, enabling faster alignment times, particularly for large angular displacements. Surprisingly, 2xHead was only faster than 1xHead in the search task, not in the alignment task. One reason could be that, at extreme angles, the high-control display gain led to overshooting, making it more difficult to perform the alignment. Combining it with other techniques for precise alignment (e.g., similar to HeadShift [62]) would alleviate this but break the absolute mapping, introducing complications. LALA was the slowest.

LALA being slower than 1xHead and 2xHead was expected, as the technique was intentionally designed and calibrated to prioritize smooth motion over speed. Scaling up LALA's velocity profiles to a competitive level — while possible — would most likely induce cybersickness [5, 29], lower stability, reduce comfort, and ultimately prevent adoption. The slower alignment times, therefore, reflect the design trade-off inherent in human-computer interaction: prioritizing users' needs over performance.

Although 2xHead achieved faster task completion times, participants preferred LALA. This highlights that objective efficiency alone does not determine user preference. While 2xHead reduces the amount of head rotation required, the high control-display gain causes a large visual-vestibular mismatch which can be uncomfortable. Additionally, it also makes precise selection more challenging. On the other hand, LALA allows participants to explore the environment primarily through natural eye movements alone. This leads to a less physical 2D exploration behaviour and is supported by the smoother rotations. This is supported by the crafted gaze control profile (1b) which prevents large accidental rotations while the user is looking around in the central area of the rotated view, only allowing extremely slow rotations that participants seem to not notice or mind. LALA also has less of a visual-vestibular mismatch compared to 2xHead because only sustained, large, eye-in-head angle causes large rotations. As sustained, large, eye-in-head angles are rare during gaze shifts involving the head, LALA induces minimal extra rotations, being similar to 1xHead. Still, several participants preferred 1xHead because its direct mapping felt the most natural. Together, these findings suggest that viewport control techniques occupy different positions along a trade-off between naturalness, efficiency, and physical effort.

These findings suggest that the three techniques occupy different positions in the design space of viewport control. Direct head-based control (1xHead) may remain most suitable for immersive and active VR experiences where natural mapping is important. Amplification approaches such as 2xHead may be preferable for tasks that require rapid orientation changes and fast navigation (such as action shooters). In contrast, gaze-driven techniques such as LALA may be particularly beneficial for prolonged viewing, slow-paced experiences (escape room games, educational content, 360° videos) or scenarios with constrained mobility (train, airplane, car).

6.2 LALA's control model for future gaze-based view control

VR is considered an emerging medium, with adoption still in its infancy [35,42]. LALA illustrates how viewport control can move beyond head rotation alone by exploring alternative interaction strategies based on gaze input, positioning it as an example of a gaze-driven control paradigm rather than a universal successor to head-based techniques.

The results show that our asymmetric omnidirectional gaze control profile (subsection 3.1) worked well for controlling viewport rotation, allowing participants to select known and search for unknown targets. Notably, an additional key benefit is that it encapsulating the complexities of eye movements, reducing the barrier of using eye movements as a control signal, simplifying downstream technique design. Prior gaze-based viewport control techniques are often tightly coupled with the visual and oculomotor systems [24,25], requiring careful calibration of individual components and making rapid prototyping difficult. In contrast, because the gaze control profile handles eye behaviour, the calculations of LALA's yaw and pitch velocities are based on relatively simple equations. We also believe our asymmetric omnidirectional gaze control profile has the potential to be useful for other gaze-based interaction techniques outside viewport control.

A potential extension of LALA is user-specific recalibration of the gaze control profile to account for differences in eye physiology and personal preferences. For example, users could roll their eyes to determine their maximum amplitude threshold (blue line in 1b), which would be incorporated into the eye-effort computation. This could also account for individual differences in the tendency to support gaze with head movement (i.e., head movers vs. non-head movers) [12, 54], a factor suggested to underlie issues in prior gaze-based viewport control techniques [25].

We believe our design choices position LALA as a promising viewport control technique. Combined with results showing strong user preference, low effort, and high subjective ratings, LALA is a strong candidate for long-term adoption in VR, particularly in accessibility and constrained contexts. We also expect LALA (and similar techniques) to become more relevant later in VR's adoption cycle, as the user base broadens to include more diverse physical abilities and as novelty diminishes, increasing demand for more ergonomic, low-effort control methods.

6.3 Limitations and Future Work

LALA is the first omnidirectional gaze-based viewport control technique for VR, enabling users to intuitively look around independently of head or torso mobility for general use. However, this study was an abstract evaluation of its feasibility and usability, conducted in a controlled environment to enable consistent comparisons rather than capture the full range of real-world VR use.

Additionally, in our study, participants spent an estimated average of approximately 40 minutes in VR, divided into six segments across an hour with breaks in between. While this reduced fatigue accumulation and supported controlled comparisons, it also limited fatigue effects that would naturally emerge during longer sessions.¹ Consequently, the results may not fully generalize to more prolonged, embodied, or immersive VR experiences.

We believe that in highly interactive or immersive applications (e.g., VR games, training, or rehabilitation), naturalness would outweigh ergonomic concerns, favoring the default 1xHead condition. Conversely, in passive, strenuous, or long-duration applications (e.g., 360° videos or virtual workspaces), minimizing effort and ergonomic strain would be prioritized, favoring LALA. Lastly, many interaction techniques use gaze for selection or manipulation. We believe LALA can coexist with these approaches due to its natural, implicit interaction style and may even enhance them by reducing required dwell times, as smooth pursuit movements [57, 60] could provide a stronger indicator of user intent, warranting further investigation.

¹Meta recommends VR experiences to be designed around a session length of 20 to 40 minutes [35]

7 CONCLUSION

In this paper, we presented Looking Around by Looking Around (LALA), an omnidirectional gaze-based viewport control technique for VR, designed to minimize physical movement and emphasize intuitiveness. We achieved this by first designing asymmetric omnidirectional control profile for the eye, then built on it to exploit tendencies for eyes to stay within comfortable regions for viewport control. We evaluated LALA in a user study (N=18) featuring two contrasting tasks: alignment towards known directions and open-ended visual search towards unknown directions. Our results showed that LALA was strongly preferred over the traditional baseline while also enabling fully hands-free interaction with minimal physical movement.

These findings underline that intuitiveness and reduced physical strain can outweigh efficiency in shaping preferences, suggesting that gaze-only viewport control is not only feasible but desirable for many VR contexts. Looking ahead, LALA offers gaze-only control for scenarios where comfort and low physical effort are essential, such as extended media viewing, collaborative work, or social VR. More broadly, LALA reimagines VR as a less physically and spatially demanding medium, expanding its accessibility and usability for prolonged or constrained use.

ACKNOWLEDGMENTS

This work was supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (Grant No. 101021229 GEMINI: Gaze and Eye Movement in Interaction, and Grant No. 101247164 EyeView360: 360° View Control by Eye Movement for Extended Reality).

REFERENCES

- [1] A. J. T. Alsmith, E. R. Ferrè, and M. R. Longo. Dissociating contributions of head and torso to spatial reference frames: The misalignment paradigm. *Consciousness and Cognition*, 53:105–114, Aug. 2017. doi: 10.1016/j.concog.2017.06.005 4
- [2] J. Ardouin, A. Lécuyer, M. Marchal, C. Riant, and E. Marchand. Flyviz: A novel display device to provide humans with 360° vision by coupling catadioptric camera with hmd. In *Proceedings of the 18th ACM Symposium on Virtual Reality Software and Technology, VRST '12*, pp. 41–44. Association for Computing Machinery, New York, NY, USA, 2012. doi: 10.1145/2407336.2407344 2
- [3] J. B. Badler and A. Canossa. Anticipatory gaze shifts during navigation in a naturalistic virtual environment. In *Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play, CHI PLAY '15*, pp. 277–283. Association for Computing Machinery, New York, NY, USA, 2015. doi: 10.1145/2793107.2793136 2
- [4] L. Bajorunaite, S. Brewster, and J. R. Williamson. Virtual reality in transit: how acceptable is vr use on public transport? In *2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pp. 432–433, 2021. doi: 10.1109/VRW52623.2021.00098 2
- [5] F. Bonato, A. Bubka, and S. Palmisano. Combined pitch and roll and cybersickness in a virtual environment. *Aviation, space, and environmental medicine*, 80:941–5, 11 2009. doi: 10.3357/ASEM.2394.2009 3, 8
- [6] J. C. Byers and S. G. Hill. Traditional and raw task load index (tlx) correlations: Are paired comparisons necessary. pp. 481,485, 1989. 6
- [7] M. Choi, D. Sakamoto, and T. Ono. Kuiper belt: Utilizing the “out-of-natural angle” region in the eye-gaze interaction for virtual reality. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems, CHI '22*, art. no. 357, 17 pp. Association for Computing Machinery, New York, NY, USA, 2022. doi: 10.1145/3491102.3517725 2
- [8] L. Cleary, K. Looney, N. Brady, and M. Fitzgerald. Inversion effects in the perception of the moving human form: A comparison of adolescents with autism spectrum disorder and typically developing adolescents. *Autism*, 18(8):943–952, Nov. 2014. Publisher: SAGE Publications Ltd. doi: 10.1177/1362361313499455 4
- [9] C. C. Collins, D. O'Meara, and A. B. Scott. Muscle tension during unrestrained human eye movements. *J Physiol*, 245(2):351–369, Feb. 1975. 3
- [10] M. Di Luca, H. Seifi, S. Egan, and M. Gonzalez-Franco. Locomotion vault: The extra mile in analyzing vr locomotion techniques. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, CHI '21*, art. no. 128, 10 pp. Association for Computing Machinery, New York,

- NY, USA, 2021. doi: [10.1145/3411764.3445319](https://doi.org/10.1145/3411764.3445319) 2
- [11] P. Eghbali, K. Väänänen, and T. Jokela. Social acceptability of virtual reality in public spaces: experiential factors and design recommendations. In *Proceedings of the 18th International Conference on Mobile and Ubiquitous Multimedia*, MUM '19, art. no. 28, 11 pp. Association for Computing Machinery, New York, NY, USA, 2019. doi: [10.1145/3365610.3365647](https://doi.org/10.1145/3365610.3365647) 2
 - [12] J. H. Fuller. Head movement propensity. *Experimental Brain Research*, 92(1):152–164, 12 1992. doi: [10.1007/BF00230391](https://doi.org/10.1007/BF00230391) 9
 - [13] K. L. Grasse and S. G. Lisberger. Analysis of a naturally occurring asymmetry in vertical smooth pursuit eye movements in a monkey. *Journal of Neurophysiology*, 67(1):164–179, 1992. doi: [10.1152/jn.1992.67.1.164](https://doi.org/10.1152/jn.1992.67.1.164) 3
 - [14] K. Grinyer and R. J. Teather. Effects of Field of View on Dynamic Out-of-View Target Search in Virtual Reality. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 139–148, Mar. 2022. ISSN: 2642-5254. doi: [10.1109/VR51125.2022.00032](https://doi.org/10.1109/VR51125.2022.00032) 5
 - [15] E. Guy, P. Punpongsonan, D. Iwai, K. Sato, and T. Boubekur. Lazynav: 3d ground navigation with non-critical body parts. In *2015 IEEE Symposium on 3D User Interfaces (3DUI)*, pp. 43–50, 2015. doi: [10.1109/3DUI.2015.7131725](https://doi.org/10.1109/3DUI.2015.7131725) 2
 - [16] J. P. Hansen, A. K. Trudslev, S. A. Harild, A. Alapetite, and K. Minakata. Providing access to vr through a wheelchair. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI EA '19, pp. 1–8. Association for Computing Machinery, New York, NY, USA, 2019. doi: [10.1145/3290607.3299048](https://doi.org/10.1145/3290607.3299048) 2
 - [17] D. Hestenes. Invariant body kinematics: I. Saccadic and compensatory eye movements. *Neural Networks*, 7(1):65–77, Jan. 1994. doi: [10.1016/0893-6080\(94\)90056-6](https://doi.org/10.1016/0893-6080(94)90056-6) 4
 - [18] C. Jay and R. Hubbard. Amplifying Head Movements with Head-Mounted Displays. *Presence: Teleoperators and Virtual Environments*, 12(3):268–276, 06 2003. doi: [10.1162/105474603765879521](https://doi.org/10.1162/105474603765879521) 2
 - [19] H. K. Kim, J. Park, Y. Choi, and M. Choe. Virtual reality sickness questionnaire (VRSQ): Motion sickness measurement index in a virtual reality environment. *Applied Ergonomics*, 69:66–73, May 2018. doi: [10.1016/j.apergo.2017.12.016](https://doi.org/10.1016/j.apergo.2017.12.016) 6
 - [20] M. J. Kobel, A. R. Wagner, and D. M. Merfeld. Impact of gravity on the perception of linear motion. *Journal of Neurophysiology*, 126(3):875–887, Sept. 2021. Publisher: American Physiological Society. doi: [10.1152/jn.00274.2021](https://doi.org/10.1152/jn.00274.2021) 3, 4
 - [21] K. Kornyo, N. Dill, M. Saenz, and R. J. Krauzlis. Canceling of pursuit and saccadic eye movements in humans and monkeys. *Journal of Neurophysiology*, 89(6):2984–2999, 2003. PMID: 12783947. doi: [10.1152/jn.00859.2002](https://doi.org/10.1152/jn.00859.2002) 2
 - [22] M. Land and B. Tatler. Looking and acting: Vision and eye movements in natural behaviour. *Looking and Acting: Vision and Eye Movements in Natural Behaviour*, pp. 1–288, 01 2012. doi: [10.1093/acprof:oso/9780198570943.001.0001](https://doi.org/10.1093/acprof:oso/9780198570943.001.0001) 2
 - [23] E. Langbehn, J. Wittig, N. Katzakis, and F. Steinicke. Turn your head half round: Vr rotation techniques for situations with physically limited turning angle. In *Proceedings of Mensch Und Computer 2019*, MuC '19, pp. 235–243. Association for Computing Machinery, New York, NY, USA, 2019. doi: [10.1145/3340764.3340778](https://doi.org/10.1145/3340764.3340778) 2
 - [24] H. S. Lee, F. Weidner, and H. Gellersen. Patterns in motion: On head- and non-head movers in vr during viewport control. In *2024 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pp. 636–641, 2024. doi: [10.1109/VRW62533.2024.00125](https://doi.org/10.1109/VRW62533.2024.00125) 9
 - [25] H. S. Lee, F. Weidner, L. Sidenmark, and H. Gellersen. Snap, pursuit and gain: Virtual reality viewport control by gaze. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, CHI '24, art. no. 375, 14 pp. Association for Computing Machinery, New York, NY, USA, 2024. doi: [10.1145/3613904.3642838](https://doi.org/10.1145/3613904.3642838) 2, 3, 6, 9
 - [26] J. Lee, H. Kim, Y. Yang, and G. J. Kim. RPG: Rotation Technique in VR Locomotion using Peripheral Gaze. *Proc. ACM Hum.-Comput. Interact.*, 8:235:1–235:19, 2024. doi: [10.1145/3655609](https://doi.org/10.1145/3655609) 2
 - [27] D. Lim, S. Shirai, M. Kobayashi, Y. Uranishi, and H. Takemura. CaliView: Continuous Viewpoint Calibration Using Dynamic Rotation Gain Control. *IEEE Transactions on Visualization and Computer Graphics*, 31(12):10942–10951, Dec. 2025. doi: [10.1109/TVCG.2025.3620419](https://doi.org/10.1109/TVCG.2025.3620419) 2
 - [28] Y.-T. Lin, Y.-C. Liao, S.-Y. Teng, Y.-J. Chung, L. Chan, and B.-Y. Chen. Outside-In: Visualizing Out-of-Sight Regions-of-Interest in a 360° Video Using Spatial Picture-in-Picture Previews. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology*, UIST '17, pp. 255–265. Association for Computing Machinery, New York, NY, USA, Oct. 2017. doi: [10.1145/3126594.3126656](https://doi.org/10.1145/3126594.3126656) 2
 - [29] W. Lo and R. H. So. Cybersickness in the presence of scene rotational movements along different axes. *Applied Ergonomics*, 32(1):1–14, 2001. doi: [10.1016/S0003-6870\(00\)00059-4](https://doi.org/10.1016/S0003-6870(00)00059-4) 8
 - [30] M. Marchal, J. Pettré, and A. Lécuyer. Joyman: a Human-Scale Joystick for Navigating in Virtual Worlds. In *IEEE Symposium on 3D User Interfaces 2011 (3DUI)*. Singapur, Singapore, Mar. 2011. 2
 - [31] N. Marek and S. Pollmann. Contextual-Cueing beyond the Initial Field of View—A Virtual Reality Experiment. *Brain Sciences*, 10(7):446, July 2020. doi: [10.3390/brainsci10070446](https://doi.org/10.3390/brainsci10070446) 5
 - [32] M. McCullough, H. Xu, J. Michelson, M. Jackoski, W. Pease, W. Cobb et al. Myo arm: swinging to explore a ve. In *Proceedings of the ACM SIGGRAPH Symposium on Applied Perception*, SAP '15, pp. 107–113. Association for Computing Machinery, New York, NY, USA, 2015. doi: [10.1145/2804408.2804416](https://doi.org/10.1145/2804408.2804416) 2
 - [33] M. McGill, A. Kehoe, E. Freeman, and S. Brewster. Expanding the bounds of seated virtual workspaces. *ACM Trans. Comput.-Hum. Interact.*, 27(3), art. no. 13, 40 pp., may 2020. doi: [10.1145/3380959](https://doi.org/10.1145/3380959) 2
 - [34] M. J. McGuffin and R. Balakrishnan. Fitts' law and expanding targets: Experimental studies and designs for user interfaces. *ACM Trans. Comput.-Hum. Interact.*, 12(4):388–422, Dec. 2005. doi: [10.1145/1121112.1121115](https://doi.org/10.1145/1121112.1121115) 5
 - [35] Meta. Building Content to Optimize User Time: 20-40 Minute “Goldilocks” Session Length on Meta Quest. 9
 - [36] C. H. Meyer, A. G. Lasker, and D. A. Robinson. The upper limit of human smooth pursuit velocity. *Vision Research*, 25(4):561–563, 1985. doi: [10.1016/0042-6989\(85\)90160-9](https://doi.org/10.1016/0042-6989(85)90160-9) 2
 - [37] M. R. Mine. Virtual environment interaction techniques. *UNC Chapel Hill CS Dept*, 1995. 2
 - [38] M. Mott, J. Tang, S. Kane, E. Cutrell, and M. Ringel Morris. “i just went into it assuming that i wouldn't be able to have the full experience”: Understanding the accessibility of virtual reality for people with limited mobility. In *Proceedings of the 22nd International ACM SIGACCESS Conference on Computers and Accessibility*, ASSETS '20, art. no. 43, 13 pp. Association for Computing Machinery, New York, NY, USA, 2020. doi: [10.1145/3373625.3416998](https://doi.org/10.1145/3373625.3416998) 2
 - [39] N. Norouzi, G. Bruder, and G. Welch. Assessing vignetting as a means to reduce vr sickness during amplified head rotations. In *Proceedings of the 15th ACM Symposium on Applied Perception*, SAP '18, art. no. 19, 8 pp. Association for Computing Machinery, New York, NY, USA, 2018. doi: [10.1145/3225153.3225162](https://doi.org/10.1145/3225153.3225162) 2
 - [40] Y. S. Pai, B. I. Outram, B. Tag, M. Isogai, D. Ochi, and K. Kunze. Gaze-sphere: Navigating 360-degree-video environments in vr using head rotation and eye gaze. In *ACM SIGGRAPH 2017 Posters*, SIGGRAPH '17, art. no. 23, 2 pp. Association for Computing Machinery, New York, NY, USA, 2017. doi: [10.1145/3102163.3102183](https://doi.org/10.1145/3102163.3102183) 2
 - [41] E. D. Ragan, S. Scerbo, F. Bacim, and D. A. Bowman. Amplified head rotation in virtual reality and the effects on 3d search, training transfer, and spatial orientation. *IEEE Transactions on Visualization and Computer Graphics*, 23(8):1880–1895, 2017. doi: [10.1109/TVCG.2016.2601607](https://doi.org/10.1109/TVCG.2016.2601607) 2
 - [42] F. Richter. Infographic: AR & VR Adoption Is Still in Its Infancy, June 2023. 9
 - [43] S. P. Sargunam, K. R. Moghadam, M. Suhail, and E. D. Ragan. Guided head rotation and amplified head rotation: Evaluating semi-natural travel and viewing techniques in virtual reality. In *2017 IEEE Virtual Reality (VR)*, pp. 19–28, March 2017. doi: [10.1109/VR.2017.7892227](https://doi.org/10.1109/VR.2017.7892227) 2
 - [44] S. P. Sargunam and E. D. Ragan. Evaluating joystick control for view rotation in virtual reality with continuous turning, discrete turning, and field-of-view reduction. In *Proceedings of the 3rd International Workshop on Interactive and Spatial Computing*, IWISC '18, pp. 74–79. Association for Computing Machinery, New York, NY, USA, 2018. doi: [10.1145/3191801.3191815](https://doi.org/10.1145/3191801.3191815) 2
 - [45] D. Sato, M. Ezoe, A. Hino, A. Shimizu, H. Mitake, and S. Hasegawa. Vibroskate: a locomotion interface with the exact haptics and kinesthesia. In *ACM SIGGRAPH 2015 Emerging Technologies*, SIGGRAPH '15, art. no. 25, 1 p. Association for Computing Machinery, New York, NY, USA, 2015. doi: [10.1145/2782782.2792502](https://doi.org/10.1145/2782782.2792502) 2
 - [46] J. Schjerlund, K. Hornbæk, and J. Bergström. Overlap: Perceiving multiple locations simultaneously to improve interaction in vr. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, CHI '22, art. no. 355, 13 pp. Association for Computing Machinery, New York, NY, USA, 2022. doi: [10.1145/3491102.3501873](https://doi.org/10.1145/3491102.3501873) 2

- [47] S. Sharmin, O. Špakov, and K.-J. Rähä. Reading on-screen text with gaze-based auto-scrolling. In *Proceedings of the 2013 Conference on Eye Tracking South Africa*, ETSA '13, pp. 24–31. Association for Computing Machinery, New York, NY, USA, 2013. doi: [10.1145/2509315.2509319](https://doi.org/10.1145/2509315.2509319) 2
- [48] L. Sidenmark and H. Gellersen. Eye, head and torso coordination during gaze shifts in virtual reality. *ACM Trans. Comput.-Hum. Interact.*, 27(1), art. no. 4, 40 pp., dec 2019. doi: [10.1145/3361218](https://doi.org/10.1145/3361218) 2, 3
- [49] P. Skalski, R. Tamborini, A. Shelton, M. Buncher, and P. Lindmark. Mapping the road to fun: Natural video game controllers, presence, and game enjoyment. *New Media & Society*, 13(2):224–242, Mar. 2011. Publisher: SAGE Publications. doi: [10.1177/1461444810370949](https://doi.org/10.1177/1461444810370949) 1
- [50] J.-P. Stauffert, F. Niebling, and M. E. Latoschik. Effects of latency jitter on simulator sickness in a search task. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 121–127, 2018. doi: [10.1109/VR.2018.8446195](https://doi.org/10.1109/VR.2018.8446195) 3
- [51] J.-P. Stauffert, F. Niebling, and M. E. Latoschik. Simultaneous run-time measurement of motion-to-photon latency and latency jitter. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 636–644, 2020. doi: [10.1109/VR46266.2020.00086](https://doi.org/10.1109/VR46266.2020.00086) 3
- [52] N. Stein, T. Watson, M. Lappe, M. Westendorf, and S. Durant. Eye and head movements in visual search in the extended field of view. *Scientific Reports*, 14(1):8907, Apr. 2024. Publisher: Nature Publishing Group. doi: [10.1038/s41598-024-59657-5](https://doi.org/10.1038/s41598-024-59657-5) 5
- [53] Q. Sun, A. Patney, L.-Y. Wei, O. Shapira, J. Lu, P. Asente et al. Towards virtual reality infinite walking: Dynamic saccadic redirection. *ACM Trans. Graph.*, 37(4), art. no. 67, 13 pp., jul 2018. doi: [10.1145/3197517.3201294](https://doi.org/10.1145/3197517.3201294) 2
- [54] Z. C. Thumser, B. S. Oommen, I. S. Kofman, and J. S. Stahl. Idiosyncratic variations in eye-head coupling observed in the laboratory also manifest during spontaneous behavior in a natural setting. *Experimental Brain Research*, 191(4):419–434, 12 2008. doi: [10.1007/s00221-008-1534-2](https://doi.org/10.1007/s00221-008-1534-2) 9
- [55] T. van Gemert, K. Hornbæk, J. Knibbe, and J. Bergström. Towards a bedder future: A study of using virtual reality while lying down. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, CHI '23, art. no. 877, 18 pp. Association for Computing Machinery, New York, NY, USA, 2023. doi: [10.1145/3544548.3580963](https://doi.org/10.1145/3544548.3580963) 2
- [56] E. Velloso and M. Carter. The emergence of eyeplay: A survey of eye interaction in games. In *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play*, CHI PLAY '16, pp. 171–185. Association for Computing Machinery, New York, NY, USA, 2016. doi: [10.1145/2967934.2968084](https://doi.org/10.1145/2967934.2968084) 2
- [57] E. Velloso, M. Carter, J. Newn, A. Esteves, C. Clarke, and H. Gellersen. Motion Correlation: Selecting Objects by Matching Their Movement. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 24(3):22:1–22:35, Apr. 2017. doi: [10.1145/3064937](https://doi.org/10.1145/3064937) 9
- [58] L. Vezzani, F. Strada, F. G. Praticò, and A. Bottino. The walkingseat: A leaning interface for locomotion in virtual environments. In *Virtual Reality and Mixed Reality: 20th EuroXR International Conference, EuroXR 2023, Rotterdam, The Netherlands, November 29 – December 1, 2023, Proceedings*, pp. 21–39. Springer-Verlag, Berlin, Heidelberg, 2023. doi: [10.1007/978-3-031-48495-7_2](https://doi.org/10.1007/978-3-031-48495-7_2) 2
- [59] M. Vidal, M.-A. Amorim, J. McIntyre, and A. Berthoz. The perception of visually presented yaw and pitch turns: Assessing the contribution of motion, static, and cognitive cues. *Perception & Psychophysics*, 68(8):1338–1350, Nov. 2006. doi: [10.3758/BF03193732](https://doi.org/10.3758/BF03193732) 3
- [60] M. Vidal, A. Bulling, and H. Gellersen. Pursuits: Spontaneous interaction with displays based on smooth pursuit eye movement and moving targets. In *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, UbiComp '13, pp. 439–448. ACM, New York, NY, USA, 2013. doi: [10.1145/2493432.2493477](https://doi.org/10.1145/2493432.2493477) 9
- [61] C. Wang, S.-H. Zhang, Y. Zhang, S. Zollmann, and S.-M. Hu. On Rotation Gains Within and Beyond Perceptual Limitations for Seated VR. *IEEE Transactions on Visualization and Computer Graphics*, 29(7):3380–3391, 2023. doi: [10.1109/TVCG.2022.3159799](https://doi.org/10.1109/TVCG.2022.3159799) 1
- [62] H. Wang, L. Sidenmark, F. Weidner, J. Newn, and H. Gellersen. HeadShift: Head Pointing with Dynamic Control-Display Gain. *ACM Trans. Comput.-Hum. Interact.*, 2024. doi: [10.1145/3689434](https://doi.org/10.1145/3689434) 8
- [63] M. Westhoven, D. Paul, and T. Alexander. Head turn scaling below the threshold of perception in immersive virtual environments. In *Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology*, VRST '16, pp. 77–86. Association for Computing Machinery, New York, NY, USA, 2016. doi: [10.1145/2993369.2993385](https://doi.org/10.1145/2993369.2993385) 2
- [64] J. O. Wobbrock, L. Findlater, D. Gergle, and J. J. Higgins. The aligned rank transform for nonparametric factorial analyses using only anova procedures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pp. 143–146. Association for Computing Machinery, New York, NY, USA, 2011. doi: [10.1145/1978942.1978963](https://doi.org/10.1145/1978942.1978963) 6
- [65] A. M. F. Wong. Listing's law: Clinical significance and implications for neural control. *Survey of Ophthalmology*, 49(6):563–575, Nov. 2004. doi: [10.1016/j.survophthal.2004.08.002](https://doi.org/10.1016/j.survophthal.2004.08.002) 4
- [66] S. Zhang, C. Wang, Y. Zhang, F.-L. Zhang, N. Pantidi, and S.-M. Hu. Velocity Guided Amplification of View Rotation for Seated VR Scene Exploration. In *2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pp. 504–505, Mar. 2021. doi: [10.1109/VRW52623.2021.00134](https://doi.org/10.1109/VRW52623.2021.00134) 2
- [67] Y. Zhang, A. Bulling, and H. Gellersen. Sideways: a gaze interface for spontaneous interaction with situated displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, pp. 851–860. Association for Computing Machinery, New York, NY, USA, 2013. doi: [10.1145/2470654.2470775](https://doi.org/10.1145/2470654.2470775) 2
- [68] Y. Zhang, M. K. Chong, J. Müller, A. Bulling, and H. Gellersen. Eye tracking for public displays in the wild. *Personal and Ubiquitous Computing*, 19(5):967–981, Aug 2015. doi: [10.1007/s00779-015-0866-8](https://doi.org/10.1007/s00779-015-0866-8) 2

APPENDIX

.1 Source Code for LALA and the study

Source code for LALA and the entire project can be found at <https://github.com/rubyleehs/LALA-Looking-Around-by-looking-Around-Public/releases/tag/LALA>

.1.1 Pseudo Code for rotation remapping

Algorithm 1 Rotation Remapping

```

1: procedure APPLYCONTROLLERSTATE(updatePhase, controller-
   State)
2:   Execute default XR controller/camera state update
3:   Rotate camera around the local (camera.right.x, 0, cam-
   era.right.z) axis by gaze-induced pitch.
4:   Rotate camera around the global up axis by gaze-induced yaw.

   ▷ The following is included for compari-
   son/replication/extensibility and not used in the current
   experiment
5:   Rotate xrRig around the local camera right axis by any addi-
   tional torso-induced pitch.
6:   Rotate xrRig around the local camera up axis by any additional
   torso-induced yaw.
7: end procedure

```

.1.2 Representative values for the gaze control profile

.2 Full Statistics

		Technique Main Effect		
		Friedman Statistic	p	W
NASA TLX	Comfort	$\chi^2(2) = 2.98$	.225	.082
	Ease of Use	$\chi^2(2) = 0.70$	.705	.019
	Controllability	$\chi^2(2) = 0.61$	.736	.017
	Mental Demand	$\chi^2(2) = 0.31$	.985	.001
	Physical Demand	$\chi^2(2) = 16.6$	<.001	.461 ***
	Temporal Demand	$\chi^2(2) = 0.58$	.750	.016
	Performance	$\chi^2(2) = 1.11$	.573	.031
VRSQ	Effort	$\chi^2(2) = 7.44$	.024	.207 *
	Frustration	$\chi^2(2) = 4.52$	.104	.126
	Oculomotor Disturbance	$\chi^2(2) = 0.76$	.683	.021
	Disorientation	$\chi^2(2) = 0.54$	.763	.015
Total Simulator Sickness		$\chi^2(2) = 0.03$	.984	.001

Table 5: Alignment Task: Friedman statistics for subjective data. Effect sizes reported as Kendall's W.

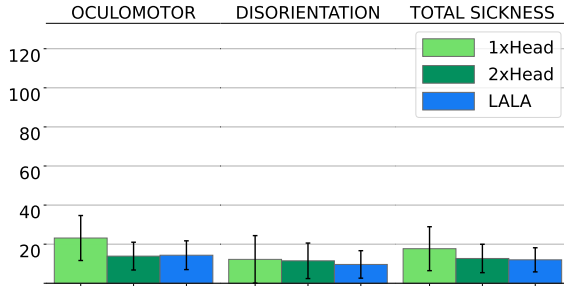


Figure 10: VRSQ by Technique for our Alignment Task. Error bars represent the 95% confidence interval. Lower values are better.

		Technique Main Effect			
		<i>F</i> value	<i>p</i>	η^2	
Alignment	Speed	$F(2,34) = 41.43$	<.001	.709	***
	Error Rate	$F(2,34) = 0.811$	.453	.046	
Cumulative Rotation of	Eye	$F(1.28,21.73) = 26.422$	<.001	.473	***
	Head	$F(2,34) = 36.58$	<.001	.683	***
	Torso	$F(2,34) = 0.818$	.450	.046	
Max Pitch of	Eye	$F(2,34) = 28.942$	<.001	.630	***
	Head	$F(1.2,20.39) = 43.012$	<.001	.569	***
	Torso	$F(2,34) = 0.359$	.700	.021	
Max Yaw of	Eye	$F(2,34) = 23.569$	<.001	.322	***
	Head	$F(2,34) = 43.932$	<.001	.721	***
	Torso	$F(2,34) = 0.266$	.768	.015	

Table 6: Search Task: ANOVA statistics. Effect sizes reflect η_g^2 when parametric assumptions were met, and η_p^2 when ART was applied.

Table 5 and Figure 10 shows the remaining full subjective statistics for our Fitts Study.

Table 6 shows the remaining full quantitative statistics for our Search Study.

Table 7 and Figure 11 shows the remaining full subjective statistics for our Search Study.

3 Search Study Results by Direction

Further analysis was done on the search study results based on the direction of the targets. The analysis showed that the results for LALA and 2xHead were largely direction-independent, whereas 1xHead exhibited significant performance degradation in the rear hemisphere.

3.1 Completion Time

Completion Time broken down by direction (see 12a) revealed a significant Technique $\times$ Direction interaction effect (see Table 8), indicating

		Technique Main Effect		
		Friedman Stat	<i>p</i>	<i>W</i>
	Comfort	$\chi^2(2) = 3.97$	.137	.110
	Ease of Use	$\chi^2(2) = 2.00$	.368	.056
	Controllability	$\chi^2(2) = 0.89$	.642	.027
NASA TLX	Mental Demand	$\chi^2(2) = 0.09$	.958	.002
	Physical Demand	$\chi^2(2) = 14.30$	<.001	.398 ***
	Temporal Demand	$\chi^2(2) = 4.15$	.126	.115
	Performance	$\chi^2(2) = 7.30$	.026	.203 *
	Effort	$\chi^2(2) = 8.48$	.014	.236 *
	Frustration	$\chi^2(2) = 7.97$	.019	.221 *
VRSQ	Oculomotor Disturbance	$\chi^2(2) = 5.67$	.059	.158
	Disorientation	$\chi^2(2) = 1.65$	.437	.046
	Total Simulator Sickness	$\chi^2(2) = 2.84$	.241	.079

Table 7: Search Task: Friedman statistics for subjective data. Effect sizes reported as Kendall's *W*.

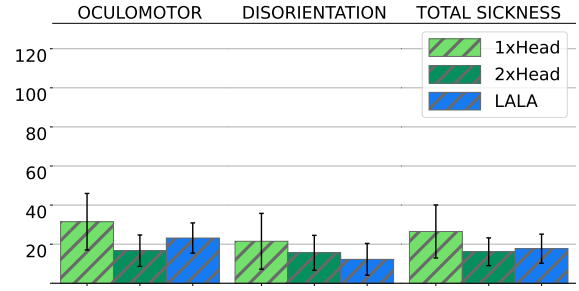


Figure 11: VRSQ by Technique for our Search Task. Error bars represent the 95% confidence interval. Lower values are better.

		Technique $\times$ Direction Interaction Effect			
		<i>F</i> value	<i>p</i>	η^2	
Alignment	Speed	$F(26,697) = 5.757$	<.001	.177	***
	Error Rate	$F(26,697) = 1.724$	.014	.060	*
Cumulative Rotation of	Eye	$F(26,697) = 5.245$	<.001	.164	***
	Head	$F(26,697) = 10.521$	<.001	.282	***
	Torso	$F(26,697) = 1.79$	.009	.063	**
Max Pitch of	Eye	$F(26,442) = 2.764$	<.001	.049	***
	Head	$F(26,442) = 3.783$	<.001	.064	***
	Torso	$F(26,697) = 1.275$	.164	.045	
Max Yaw of	Eye	$F(26,442) = 3.074$	<.001	.052	***
	Head	$F(26,697) = 3.874$	<.001	.126	***
	Torso	$F(26,697) = 4.201$	<.001	.135	***

Table 8: Search Task by Direction: ANOVA statistics. Effect sizes (η^2) reflect η_g^2 when parametric assumptions were met, and η_p^2 when ART was applied. Non-integer degrees of freedom are due to Greenhouse–Geisser corrections.

that Technique affected Completion Time differently at different Directions. Post-hoc pairwise analysis revealed that while LALA and 2xHead had no within-technique differences between Directions (all $p = 1$), 1xHead had significantly slower alignment times for targets in the Back Direction, compared to those to the sides or in the frontal hemisphere (all $p < .021$). Meanwhile, in the Back and Up Directions, 2xHead was significantly faster than both 1xHead and LALA (all $p < .002$), while in the Front direction, 1xHead outperformed LALA ($p = 0.02$).

3.2 Error Rate

Error Rate broken down by Direction revealed a significant Technique $\times$ Direction interaction (see Table 8). Post-hoc pairwise comparisons, however, showed no significant differences between individual conditions, indicating that Error Rates were similar across the board.

3.3 Max Eye, Head, and Torso Rotations

Analysis of Max Eye Pitch broken down by Direction (see 14a) revealed a significant Technique $\times$ Direction interaction effect (see Table 8), indicating that Technique affected Max Eye Pitch differently at different Directions. Post-hoc within-technique analysis revealed that only 2xHead had no significant differences in Max Eye Pitch across directions (all $p = 1$).

Analysis of Max Eye Yaw broken down by Direction (see 14a) revealed a significant Technique $\times$ Direction interaction effect (see Table 8), indicating that Technique affected Max Eye Yaw differently at different Directions. Post-hoc within-technique analysis revealed that while both LALA and 2xHead had no significant differences in Max Eye Pitch across directions (all $p = 1$), 1xHead had significantly higher Max Eye Yaw for Back compared to all directions in the frontal hemisphere and sides (all $p < .039$).

Analysis of Max Head Pitch broken down by Direction (see Figure 13) revealed a significant Technique $\times$ Direction interaction effect

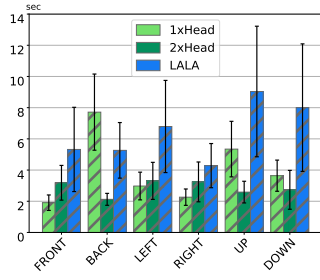


Figure 12: Completion Time By Direction - only cardinal directions are shown for conciseness. Error bars represent the 95% confidence interval. Lower values are better.

(see Table 8), indicating that Technique affected Max Head Pitch differently at different Directions. Post-hoc analysis revealed that both LALA and 2xHead had no significant differences in Max Head Pitch across directions (all $p = 1$), and had significantly lower Max Head Pitch than 1xHead in all Directions.

Analysis of Max Head Yaw broken down by Direction (see Figure 13) revealed a significant Technique $\times$ Direction interaction effect (see Table 8), indicating that Technique affected Max Head Yaw differently at different Directions. Post-hoc analysis revealed that both LALA and 2xHead had no significant differences in Max Head Pitch across directions (all $p = 1$). Meanwhile, 1xHead had significantly higher Max Head Yaw at the Back compared to the Front ($p < .001$), had significantly higher Max Head Yaw in all directions compared to LALA (all $p < .006$), and for all directions in the rear hemisphere plus sides compared to 2xHead (all $p < .002$).

Analysis of Max Torso Pitch by Direction showed no significant Technique $\times$ Direction interaction. However, there was a significant interaction for Max Torso Yaw by Direction (see Table 8). Post-hoc pairwise comparisons for Max Torso Yaw revealed significant differences only for 1xHead, with the Back direction showing higher values than all other directions.

3.4 Cumulative Eye, Head, and Torso Rotations

Cumulative Eye Rotation broken down by direction (see 14a) revealed a significant Technique $\times$ Direction interaction effect (see Table 8), indicating that Technique affected Cumulative Eye Rotation differently at different Directions. Post-hoc within-technique pairwise analysis revealed that while LALA and 2xHead had no significant differences between Directions (all $p = 1$), 1xHead had significantly more Cumulative Eye Rotation for targets in the upper-rear semihemisphere — including the Back direction — compared to those to the sides or in the frontal hemisphere (all $p < .032$). Several between-technique, within-direction differences were also observed. Notably, 1xHead showed less Cumulative Eye Rotation than LALA for the Front direction ($p=0.18$), while 2xHead had significantly lower rotation than both 1xHead and LALA in the Up direction and most of the directions in the rear hemisphere (all $p < .017$). Additionally, both 1xHead and 2xHead had significantly less Cumulative Eye Rotation than LALA for the Front-Up-Right direction (all $p < .011$).

Cumulative Head Rotation broken down by direction (see 14b) revealed a significant Technique $\times$ Direction interaction effect (see Table 8), indicating that Technique affected Cumulative Head Rotation differently at different Directions. Post-hoc within-technique pairwise analysis revealed that while LALA and 2xHead had no significant differences between Directions (all $p = 1$), 1xHead had significantly more Cumulative Head Rotation for targets in the upper-rear semihemisphere — including the Back direction — compared to those to the sides or in the frontal hemisphere (all $p < .017$). Several between-technique, within-direction differences were also observed. When compared to LALA, 1xHead showed consistently higher head rotation across all directions (all $p < .012$) except Front-Up-Right ($p = 1$). Additionally, for the Up, Down, and all directions in the rear hemisphere, 1xHead

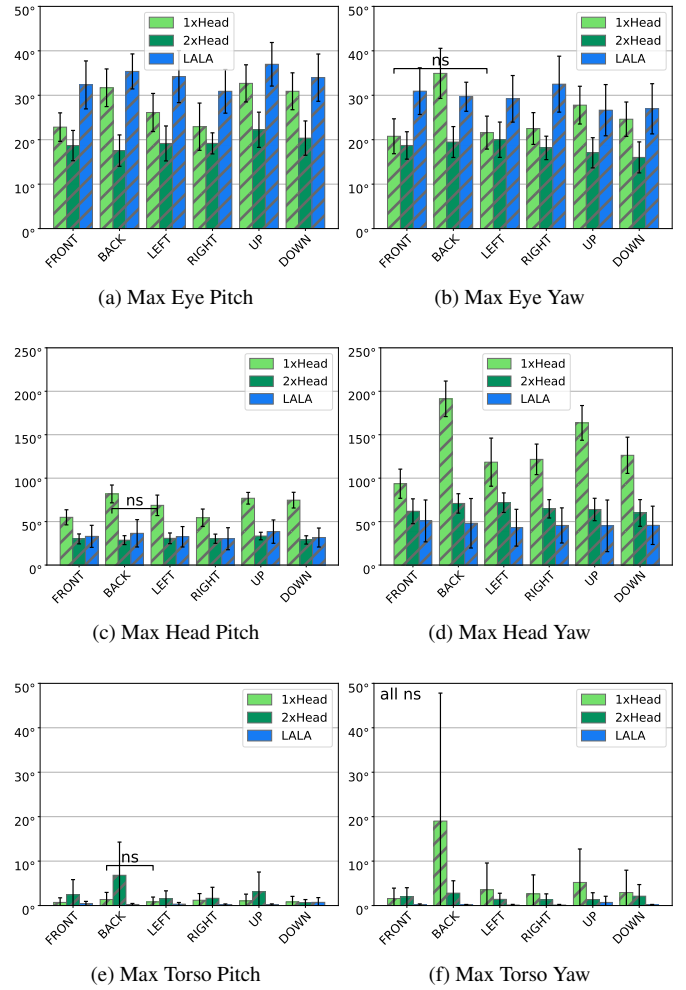


Figure 13: Max Eye (a), Head (c), and Torso Pitch (e) by Direction and Max Eye (b), Head (d), and Torso Yaw (f) by Direction - only cardinal directions are shown for conciseness. Error bars represent the 95% confidence interval. Lower values are better.

produced significantly more Cumulative Head Rotation than 2xHead (all $p < .001$).

Cumulative Torso Rotation broken down by Direction revealed a significant Technique $\times$ Direction interaction (see Table 8). Post-hoc pairwise comparisons, however, showed no significant differences between individual conditions (all $p = 1$), indicating that Cumulative Torso Rotations were similar across the board. The overall mean Cumulative Torso Rotation across all conditions was $M = 7.116^\circ$ (95% CI = $\pm 1.727^\circ$).

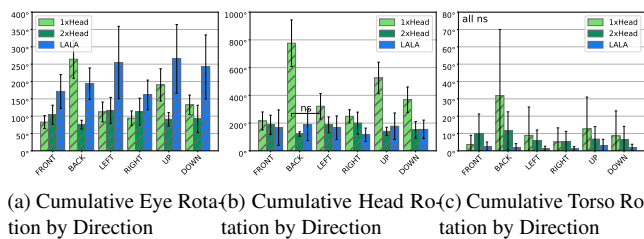


Figure 14: Cumulative (a) Eye, (b) Head, and (c) Torso Rotation by Direction - only cardinal directions are shown for conciseness. Error bars represent the 95% confidence interval. Lower values are better.