Qeios

Research Article

Gaze-Hand Steering for Travel and Multitasking in Virtual Environments

Mona Zavichi[1], André Santos[1], Catarina Moreira[2], Anderson Maciel[1], Joaquim Jorge[1]

1. University of Lisbon, Lisbon, Portugal; 2. University of Technology Sydney, Australia

As head-mounted displays (HMDs) with eye-tracking become increasingly accessible, the need for effective gaze-based interfaces in virtual reality (VR) grows. Traditional gaze- or hand-based navigation often limits user precision or impairs free viewing, making multitasking difficult. We present a gaze-hand steering technique that combines eye-tracking with hand-pointing: users steer only when gaze aligns with a hand-defined target, reducing unintended actions and enabling free look. Speed is controlled via either a joystick or a waist-level speed circle. We evaluated our method in a user study (N=20) across multitasking and single-task scenarios, comparing it to a similar technique. Results show that gaze-hand steering maintains performance and enhances user comfort and spatial awareness during multitasking. Our findings support the use of gaze-hand steering in gaze-dominant VR applications requiring precision and simultaneous interaction. Our method significantly improves VR navigation in gaze-dominant, multitasking-intensive applications, supporting immersion and efficient control.

Corresponding author: Joaquim Jorge, jorgej@tecnico.ulisboa.pt

1. Introduction

With affordable HMDs now incorporating eye tracking, effective gaze-dominant interfaces have become crucial for VR environments. Multitasking in VR poses challenges, as traditional gaze- or hand-based controls often compromise precision or limit the freedom to look around. We present a novel gaze-hand steering technique that addresses these issues with eye tracking and hand pointing.

In VR, navigation techniques are essential for effective user interaction as they enable the user to reach the location of objects that will be selected and manipulated within the virtual environment. Wayfinding is the element of navigation that helps users understand their position and plan their paths, often utilizing visual cues or maps. Travel, the motor aspect of navigation, uses specific techniques to allow users to determine the speed and direction with which their viewpoint will move through the virtual space. Although selection and manipulation tasks are frequently separated from navigation, the user concludes navigation before starting

manipulation, it is not rare that the application demands selection and manipulation during locomotion, another word for travel. Selection, for example, enables users to choose objects or destinations using gaze, gestures, or controllers during motion. These elements enable users to explore and interact seamlessly within VR, enhancing immersion and spatial awareness.

Room size often limits the physical movement a user can make when applying locomotion in VR. Techniques such as joystick-based controls, gaze-directed movement, and walking-in-place have attempted to address these issues and enhance immersion without requiring large physical spaces. One relatively successful technique is the Magic Carpet[1], which applies a flying carpet metaphor to separate direction and speed controls into two phases. It works with speed control methods such as a joystick, speed ring, and walking-in-place, combined with head-gaze as a pointing technique for direction control. This accommodates the naturalness of having a ground reference and the expeditiousness of flying. However, techniques relying solely on gaze direction for control have limitations. While gaze-oriented movement can provide precise and effortless steering[2], it can also restrict users' ability to explore the environment freely, as gaze-based navigation often leads to unintentional actions, a phenomenon known as the "Midas Touch" problem[3]. This limitation prevents users from looking around without unintentionally changing direction, impacting their overall navigation freedom.

On the other hand, while hand-pointing techniques offer precision, they prevent users from simultaneously using their hands for secondary tasks like selection or manipulation, which reduces their practicality in multitasking scenarios. Adding a trigger to switch between navigation-pointing and selection- or manipulation-pointing can make interactions cumbersome and increase the likelihood of errors.

In response to these limitations, our study proposes a technique combining eye-gaze and hand-pointing to enhance speed control and steering precision while supporting multitasking and offering a free look capability. This combined technique builds upon the strengths of gaze and hand-directed methods[4][5][6][7][8], enabling more seamless navigation that retains performance even during complex or multitasking activities.

Hence, we investigate the hypothesis that combining the gaze control techniques with pointing techniques can significantly improve speed and steering control. This potentially leads to the performance not being affected when exploring and performing other tasks while navigating, and leads to more accurate travel in the virtual environment, as gaze-oriented steering is considered to be the most efficient when more complex movements are involved[9][10].

We thus focus on answering the following research question: **How does combining gaze-directed steering with hand-pointing techniques allow exploration and multitasking without compromising performance?**

In the quest for responses, we conducted an experimental study with users to assess the user experience and travel efficiency achieved with the proposed technique. We also compare the performance of our method with two different speed control techniques—joystick and speed circle—and introduce new tasks to explore the free-look capability and selection interactions, supporting multitasking.

2. Related Work

Travel is a critical task in VR, encompassing directional control and movement through virtual spaces. Steering-based methods are commonly used for VR navigation, where users can control the direction of movement as if steering a vehicle. However, traditional approaches that rely solely on gaze or hand input often suffer from limitations, such as physical fatigue or imprecise control, which recent studies aim to address by integrating gaze for quick directional indications and hand movements for refined speed and directional adjustments[4]. Despite these innovations, there is still a need for techniques that allow users to navigate while multitasking without compromising control precision or causing physical discomfort. Previous works have focused primarily on either selection or navigation but not adequately on how users can multitask during flight locomotion.

Medeiros et al.[1] studied flight locomotion by dividing the process into direction indication and speed control, utilizing a floor proxy with full-body representation. In the direction study, three techniques—Elevator+Steering, Gaze-Oriented Steering, and Hand Steering—were evaluated with 18 participants. The participants provided feedback through questionnaires, assessing user preferences, comfort, embodiment, and immersion. The Elevator+Steering technique enabled horizontal navigation through gaze and vertical navigation via buttons. Gaze-oriented and Hand-oriented techniques offered direction control based on head rotation and dominant hand movement, respectively. For the Speed study, they explored three techniques: joystick, speed circle, and WIP. The joystick controlled speed with an analog stick. The Speed Circle approach used the body as an analog stick, derived from a virtual circle metaphor[11][12]. The last technique, adapted from[13], used knee movement for a maximum velocity of 5 m/s.

The experiment involved navigating through a city scene from Unity3D Asset Store, with rings indicating the direction of the movement. The first experiment tested steering control with speed controlled by a button, and the second focused on speed control using a hand technique for direction. The first experiment showed that gaze and hand-oriented steering techniques had close results, with gaze showing shorter path length and total time and the hand technique with fewer collisions. In the second experiment, only the hand technique was used for direction control due to limited visual exploration with the gaze method.

Lai et al. investigated gaze-directed steering on user comfort in virtual reality environments[14]. Previous research has been inconclusive about which technique is best. Some found Gaze-directed better than Hand-directed[15][9], while others found the opposite[16][17]. They compared gaze, hand, and torso-directed steering, finding that gaze-directed methods significantly reduced simulator sickness symptoms like nausea and dizziness.[18] explored gaze in Virtual Reality interactions, testing if gaze-oriented steering provides benefits over existing techniques. They argued that hand input in VR interactions is often redundant with gaze information, potentially reducing arm fatigue and improving interactions. They focused on terrain navigation, identifying speed control with an analog joystick as a main problem, and offering fixed speeds. Another issue was the inability to orbit around a region of interest. They proposed using gaze to lock areas of interest and orbit using a joystick or tablet gestures.

Lai et al.[19] replicated a dual-task methodology to compare steering-based techniques with target-based ones, finding steering techniques afford greater spatial awareness due to continuous motion. However, steering techniques may increase cybersickness[20]. VR developers are recommended to provide Gaze-directed steering as an intermediate option between novice-friendly teleport and expert-friendly Hand-directed steering.

Zielasko et al.[21] mentioned the advantages of seated VR for ergonomics and long-term usage. They found torso and gaze steering effective for direction while leaning caused fatigue and cybersickness. However, gaze/view-directed steering has issues inspecting the environment independently of movement direction, commonly called the "Midas Touch" problem[3], where gaze interactions trigger unintended actions.[22] addressed this by introducing head-tilt motions for independent "free-look" control.

Combining gaze with other input modalities, particularly hand gestures, has emerged as a solution to overcome the limitations of gaze-only interaction. Techniques that utilize gaze for initial selection and confirm actions with a secondary hand gesture have effectively reduced the "Midas Touch" effect by decoupling selection from confirmation inputs[7][8]. [7] introduced Gaze-Hand Alignment, where the alignment of gaze and hand inputs triggers selection, offering a seamless and efficient approach to interaction tasks like menu selection in AR. This technique leverages the natural coordination of gaze and hand, allowing users to pre-select a target with their gaze and confirm it by aligning their hand, thereby reducing physical strain and increasing interaction speed[6]. Recent advancements, such as the study by[5], extend this approach by combining gaze rays with controller rays, effectively disambiguating target selection to reduce selection time and enhance accuracy. Similarly, the study by[4] combines gaze for initial direction setting with hand-based adjustments, allowing users to control navigation speed and direction without unwanted activations. Also, HMDs can impact spatial awareness and task performance, especially when body representation and perspective are manipulated[23].

The integration of gaze and hand inputs has improved significantly over single-modality systems, providing more natural, efficient, and accurate interaction in VR environments. Despite these advancements, multitasking within navigation, particularly during flight locomotion, needs further development. Our approach extends gaze–hand-directed steering techniques explored in previous studies, e.g.[5],[5],[4],[4],[1],[1], introducing a mechanism that supports flight locomotion and enables users to navigate efficiently without losing orientation while multitasking.

3. Our Gaze-hand Steering Mechanism
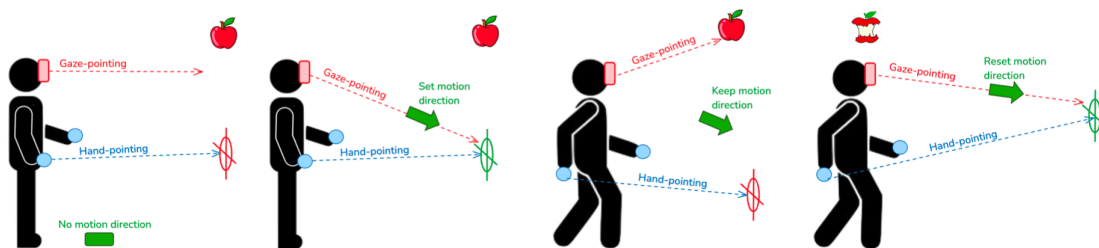
Figure 1. Travel direction is set only when gaze and hand pointing directions coincide. Travel continues in the direction set while hands and eyes can do other tasks, such as eating apples or popping balloons.

Hand pointing is an affordable and handy solution to indicate travel direction. However, currently, many HMDs incorporate eye-tracking capability, and several eye-dominant techniques are prevailing. One piece of evidence is that industry standards are adopting eye-gaze as the default dominant paradigm, as is the case with the high-end consumer device Apple Vision Pro[1]. A major constraint in eye-gaze-directed techniques, however, is the free-look or "Midas Touch" problem[3][22], which makes it difficult for users to look around without affecting movement direction. To address this, we implemented a mechanism that changes between free-look (exploring the environment with gaze without changing movement direction) and directed look (gaze changes movement direction) that allows multitasking while traveling. This mechanism combines eye-tracking gaze techniques and hand-steering, combining the two vectors to decide the direction vector. Inspired by the gaze as a helping hand study[18], we combine eye-gaze and hand pointing to lock direction and allow users to explore while maintaining movement direction. This approach uses an invisible string metaphor, where the direction locks with the interception of the eye and hand vectors, enabling users to move while looking around.

3.1. Pointing direction

The direction is null at the start, and no travel is possible. The dominant hand trigger activates the display of a target object 1 m wide at a 5 m distance along the hand-pointing direction. Then, a ray is cast in the direction of the eye-gaze, and if this ray intercepts the hand-driven target object, a new direction is set for travel (see Figure 1). The chosen direction is that of the eye gaze, providing a more egocentric definition of the direction to follow. Both eye-gaze and hands can do other tasks such as selecting or manipulating objects at any time. A new direction will only be set when the two directions align while the dominant hand trigger is pressed.

3.2. Controlling speed

Whenever a direction is set, travel speed will determine the quantity of motion in that direction. The literature has plenty of solutions. We tested two simple techniques to control speed: the joystick and the speed circle. Joystick uses the analog thumbstick to smoothly set a velocity forward or backward. A full push on the stick will set the maximum speed, or 5 m/s, which is the same used by the Magic Carpet (see Figure 6). Speed circle, in turn, uses the user's body as a joystick, i.e., leaning forward or backward sets a speed proportional to the amount of leaning (see Figure 7). Returning to the initial neutral position sets the speed to zero.

3.3. Multitasking

As we intend to use multitasking, the interface includes other simple actions. Eye gaze is the dominant input for selection. So, while it is used to select the hand-defined target for travel direction, it is also used to select other objects. One example is to point and pop balloons. This action is triggered when the eye gaze hits a balloon and the non-dominant hand trigger is pulled. To accurately shoot a target, users needed to look away from the directional scope and explore the environment to locate the targets. Upon identifying a target, they had to press the trigger on the controller in their non-dominant hand only when their line of sight intersected with the target object. This setup allowed users to multitask, managing navigation and target acquisition simultaneously.
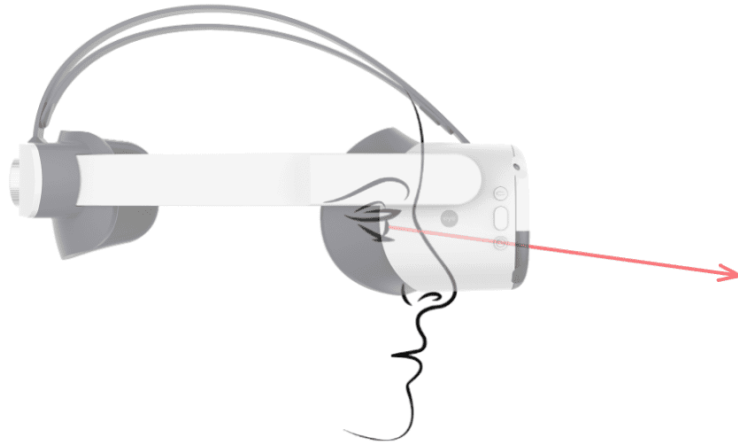
Figure 2. Eye gaze points objects for selection.
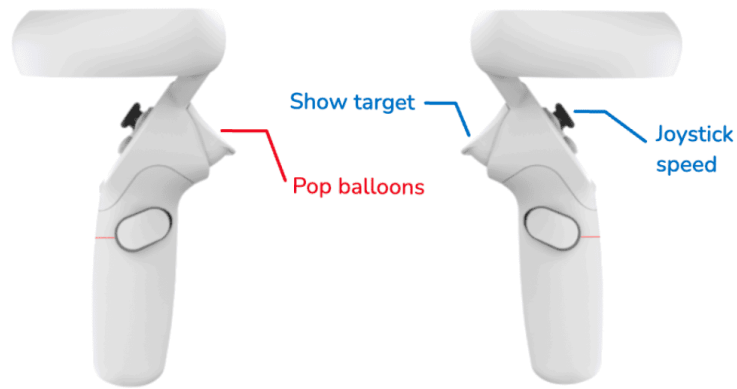
Figure 3. Left trigger confirms selection. Right trigger displays the target for travel.

Figures 2 and 3 summarize the controllers and actions.

3.4. Implementation

We implemented the above techniques for the PICO Neo 3 Pro EYE VR Headset, which supports eye-tracking, and the two included hand controllers. The eye-tracking API used with that device is the Tobii Extended Reality toolkit. Nevertheless, any similar contemporary VR kit, such as the MetaQuest Pro, could alternatively be used.

We used Unity version 2021.3 for the virtual environment and retained most assets and design elements from Magic Carpet[1] to enable direct comparative evaluation. One difference is that we did not include the walking-

in-place technique due to subpar performance results, and thus did not use the trackers they used. Another is that the Magic Carpet was based on a CAVE, whereas we used an HMD. Therefore, we also refined the Speed Circle technique to HMD environments. Our redesign sought to minimize user movement and reduce confusion regarding the circle's center. To this end, the speed circle was positioned at waist level and resized to dimensions of 0.6 x 0.6 meters, aligning it with the user's physical space and preventing abrupt accelerations during sharp turns. This enhanced stability ensures effectiveness and averts unintended movements by establishing a maximum speed threshold of 5 m/s and improving access to a null velocity zone. We also added graphical elements, such as redesigned semicircle colors with positive and negative symbols in each sector to indicate the areas for increasing or decreasing velocity. Leaning or moving towards the designated green area, marked with a plus sign, increases velocity, while moving towards the red semicircle, marked with a negative sign, decreases velocity.

As for the experimental tasks, which will be further discussed in Sec. 4, we adjusted visibility to keep all rings visible, minimizing wayfinding difficulties to focus on the eye-tracking performance.

4. Evaluation

We designed and conducted a user study to assess and characterize the effectiveness and efficiency of our steering locomotion method. We seek to establish the feasibility of using eye-tracking steering and hand-pointing gestures for navigation, enabling users to freely explore the environment. Later, we also compare our results with those from the Magic Carpet technique[1].

We anticipate observing effects in several interaction metrics. So, we measured total completion times for tasks, the occurrence of collisions, path lengths traveled, scores in the presence and workload assessments, and cybersickness. Furthermore, we anticipate introducing free look capabilities, facilitated by eye-tracking technology, that should not adversely impact flying or completion times. We expect flying and idle times to remain consistent across tasks, indicating that incorporating free-look functionality does not hinder overall performance. Additionally, we anticipate that users will perform simultaneous tasks involving traversing rings and destroying targets without undermining performance and experience when compared to data in the literature.

4.1. Environment

The virtual environment for this study is the same as that of the Magic Carpet study to facilitate direct comparison. The environment was based on a city scene with a path of rings through which the user must fly, as seen in Figs. 4 and 5, obtained from the Unity3D Asset Store. This scene was modified to remove visual

clutter and present a smoother path for testing. The interaction was implemented with the help of Unity XR's interaction kit, a framework for implementing virtual and augmented reality functionality in Unity projects.
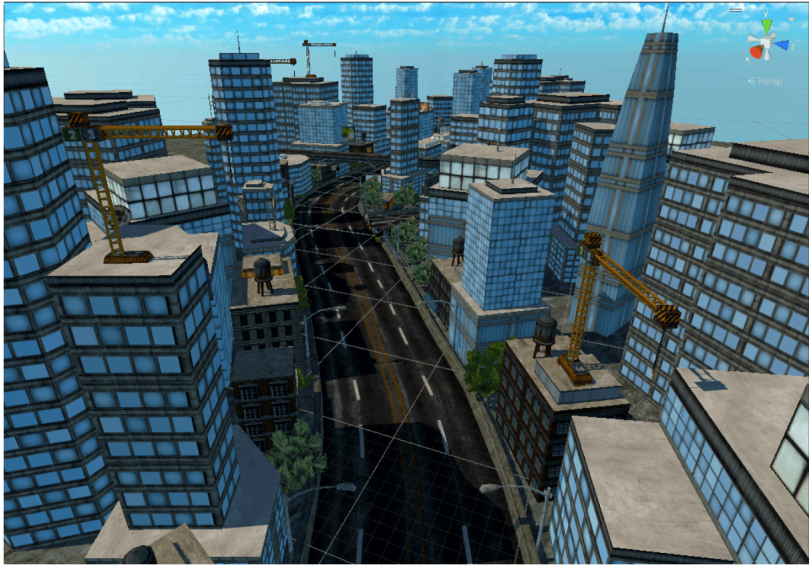

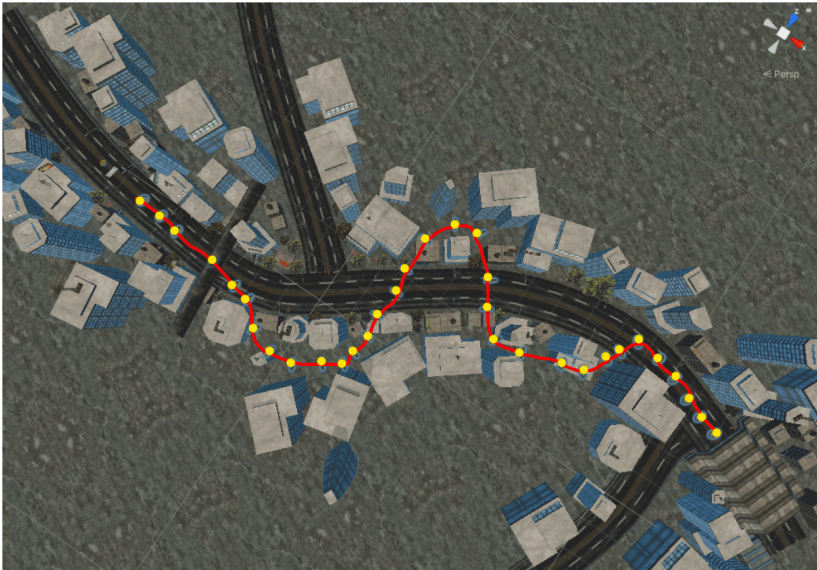**Figure 4.** Perspective of the virtual environment.

Figure 5. Path used in the experiment (350 m). Yellow dots indicate the rings placement.

4.2. Conditions

We designed the experiment to evaluate the impact of gaze-hand steering in immersive environments by assigning participants specific tasks under varying conditions—each task condition aimed at different aspects of user interaction to assess steering and targeting capabilities more effectively. The following subsections describe the functions that participants performed during the study.

4.2.1. Tasks

There are two task conditions. One is a **single-task** race where participants fly through **Rings** scattered across the map, with obstacles between the rings to test effective steering without colliding. Rings make a sound when crossed to indicate the path.

The other task condition is **multitask**. It involves a **Target** practice challenge in which users navigate through rings and shoot balloon targets placed at several locations along the path. Balloons are popped by looking at them and pressing the non-dominant hand trigger. This design focused on assessing the effectiveness of gaze-hand steering in isolation, providing insights into its performance characteristics. Comparisons with traditional eye-gaze steering are discussed using published data from the Magic Carpet experiment[1].

Figure 6. User using joystick

Figure 7. User using Speed Circle

4.2.2. Interface conditions

We implemented our gaze-hand steering as the only interface for travel direction. However, we use two conditions for speed control. Gaze-hand steering is coupled with one of the speed control interactions, making two interface conditions: Gaze-hand steering + joystick and Gaze-hand steering + speed circle.

4.2.3. Experimental conditions

Our user study thus exposed participants to four different conditions:

- Gaze-hand steering – Joystick – Rings
- Gaze-hand steering – Joystick – Targets
- Gaze-hand steering – Speed circle – Rings
- Gaze-hand steering – Speed circle – Targets

The four conditions use Gaze-hand steering to control direction. Arguably, having only one direction control condition in the protocol limits the results to findings that characterize the technique in isolation. However, we did not include other direction techniques to maintain a manageable session length and minimize participant fatigue. It would also be interesting to see how participants perform if simple eye-gaze steering is used compared to how they perform with our technique. We compare the two methods in Sec. 5.4 using published data from the Magic Carpet experiment[1].

4.3. Apparatus

To control for extraneous factors and ensure a consistent environment, all trials were conducted in a virtual reality laboratory with controlled lighting and sound levels to minimize distractions. Participants were equipped with a Head Mount Display (HMD) and calibrated hand controllers at the beginning of each session. Calibration ensured precise eye tracking, and participants were given a short training session to familiarize themselves with the gaze-hand alignment technique and verify the alignment of the equipment.

We collected both objective and subjective measures. The system logged performance measures in four aspects: task speed (time to complete), collisions (with obstacles), path length, and movement time (percentage of time passed moving instead of idling). We also administered three utility and usability measurement scales: Slater-Usoh-Steed (SUS) presence questionnaire[24], the Simulator Sickness Questionnaire (SSQ)[25], and the NASA Task Load Index (NASA-TLX)[26].

Some of these instruments were explicitly chosen to allow comparison with the results of the Magic Carpet study.

4.4. Procedure

The experimental procedure followed a within-subjects design, with participants experiencing the four conditions once. The interface condition presentation order was counterbalanced, and the task condition was fixed with the single-task Ring being performed first and the multitask Target being performed second.

After a brief explanation, participants signed an informed consent form and were told they could stop at any time if they felt discomfort, cybersickness, or other reasons. Then, they filled out a pre-test profile form. After that, they were introduced to the VR headset, the eye-tracking calibration was performed, and they were guided to the marked spot on the floor and allowed to perform the test task, where they were free to explore and practice. When they were comfortable, they started the experimental tasks.

After exposure to all four conditions, participants completed the post-test usability and experience questionnaires of Sect. 4.3 and a list of subjective questions about the system. We thanked the participants, and

they were dismissed.

The experience lasted approximately 40 minutes per participant, including the 15-minute briefing, training, and questionnaires.

4.5. Participants

Twenty participants, aged 17 to 35 ($\mu \approx 23.95$, $\sigma \approx 5.36$), volunteered for the study. The majority had at least a bachelor's degree. Despite most users having some VR experience, most were inexperienced, having used VR less than five times. Additionally, two-thirds of the participants never or sometimes experienced dizziness or nausea using VR, while the other third never used it.

5. Results

We present our results in this section according to the evidence found. We start by analyzing usability from subjective measurements and then go into the objective data collected. In the objective data part, we first look into our user performance data in isolation, developing along different tasks and interface conditions. Then, we compare our data with the results from the Magic Carpet[1] in Sec. 5.4.

5.1. Subjective Scores

This section evaluates the integrity of the developed system using SUS-presence, SSQ, and NASA-TLX questionnaires. It includes user impressions of each speed technique to compare against each other and understand how our steering approach performs overall. Statistical results are examined. Then, we present the users' responses to additional general questions regarding the experience.

A summary of the subjective instruments results from 20 participants is shown in Table 1, providing a Mean (μ), Standard Deviation (σ), and 95% Confidence Interval (CI).

First, the average SUS-presence score is 15.54 out of 20 total points (77.7%). Then, we observed that none of the participants felt severe sickness symptoms. Only four marked at least two of the 16 symptoms as moderate. The SSQ total score at the end of the participation is 21.25, but the nausea factor is lower, 13.87 on average. Disorientation was scored higher, at 30.37. For task load, the total average NASA-TLX score falls below 29 on a scale of 100, indicating a medium workload category.

Questionnaire	Metric	μ	σ	CI
SUS	Presence SUS Score	15.54	3.55	[14.06, 17.03]
SSQ	Cyber Sickness Score	21.25	18.97	[13.31, 29.18]
NASA-TLX	Mental Demand	2.41	1.05	[1.94, 2.87]
	Physical Demand	1.55	0.73	[1.22, 1.87]
	Temporal Demand	2.91	1.34	[2.31, 3.50]
	Performance	2.23	0.81	[1.86, 2.58]
	Effort	1.54	0.73	[1.22, 1.87]
	Frustration	1.45	0.73	[1.12, 1.78]
	TOTAL SCORE	28.91	6.89	[25.86, 31.96]

Table 1. Summary of the subjective metrics

5.2. Additional subjective Impressions

Table 2 summarizes the participants' impressions of using the system on a five-point Likert scale. The responses indicate that walking inside the virtual circle is significantly easier (Wilcoxon = 28.5, p=0.020) with the joystick (avg= 4.36, SD= 0.90) than with the speed circle (avg= 3.55, SD= 1.01). Controlling speed was significantly easier (Wilcoxon= 32.0, p=0.033) with the joystick (avg= 4.13, SD= 0.83) compared to the speed circle (avg= 3.31, SD= 1.13).

Moving around the virtual environment was also easier with the joystick (Wilcoxon = 12.5, $p = 0.034$). Users reported a higher sense of agency with the joystick (average = 4.32, SD = 3.45) compared to the speed circle (average = 3.45, SD = 1.14) (Wilcoxon = 0.0, $p = 0.003$). Users also felt safer inside the circle when using the joystick (Wilcoxon = 9.0, $p = 0.046$). These results collectively suggest that the joystick provides a more familiar user experience as is commonly used in video games. However, some users preferred the speed circle for its more immersive experience, allowing direct interaction through body gestures. Notably, some users achieved better results with the speed circle. Despite this, the joystick was generally preferred for speed tasks. Some users found it easier to maintain movement with the speed circle during target practice, as the joystick required more frequent stops for slow movements. However, the speed circle was more prone to imbalance issues, requiring constant position adjustments when changing direction. One user commented, *"I felt that with*

the speed circle I was more in control, but with the stick it was more immediate. In the speed circle, I was much more likely to overbalance and fall, whereas with the joystick I wasn't".

No significant differences were found between the two techniques regarding fatigue or fear of heights. Similarly, there were no significant differences in body ownership or sense of self-location. Both methods were perceived as equally effective in maintaining user immersion and safety.

Question	Joystick	Speed
Q1. It was easy to walk inside the circle.	5 (1)	4 (1)
Q2. It was easy to indicate the direction of movement.	4.5 (1)	4 (1.75)
Q3. It was easy to control the speed of movement.	4 (1)	3.5 (1.75)
Q4. It was easy to move around the VE.	4 (1)	4 (1)
Q5. It was easy to reach the rings.	4 (0.75)	3.5 (1)
Q6. It was easy to avoid obstacles.	4 (1.75)	3.5 (1.75)
Q7. It was easy to coordinate movements.	4 (1.75)	3 (1)
Q8. I felt safe inside the circle.	5 (1)	4 (1)
Q9. I felt a fear of heights.	0 (0.75)	0 (1)
Q10. I felt fatigue.	0 (0)	0 (0)
Q11. I felt a sense of agency.	4 (1)	4 (1)
Q12. I felt body ownership.	4 (1)	4 (1.75)
Q13. I felt a sense of self-location.	4 (1)	4 (1)

Table 2. Subjective impressions collected and presented as Median (Interquartile Range) Values

5.3. Objective measurements

This section examines task performance across different techniques and tasks. The analysis includes statistics with a 95% confidence interval for each metric to compare techniques, tasks, and previous findings. Both objective metrics and subjective scores were analyzed to evaluate differences between techniques. The normality of the samples was assessed using the Shapiro–Wilk test. Although most conditions followed a normal distribution, the Targets_SpeedCircle condition significantly deviated from normality ($p = 0.001$). To

ensure consistency and robustness, we proceeded with the Kruskal–Wallis nonparametric test for all comparisons to identify significant effects. Post hoc pairwise comparisons were performed using Dunn's test with Bonferroni correction to determine specific differences between conditions. The main goal was to observe to what extent our technique supports a secondary task without penalizing performance outcomes.
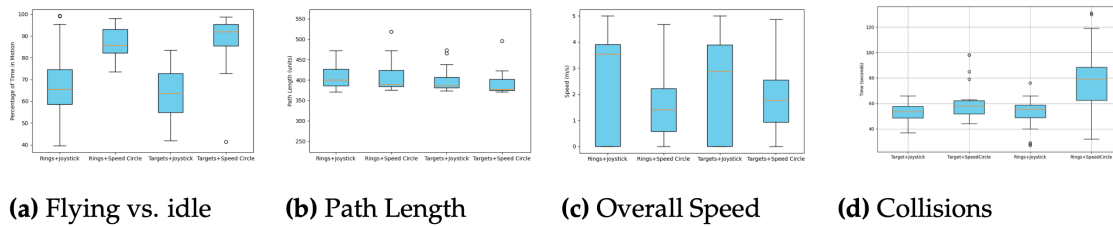
Time

Task completion time shows that the Joystick method is faster in both the Ring task ($t(38) = -4.172, p{<}0.01$) and the Target practice ($t(38) = -2.314, p = 0.026$). No significant differences were found between tasks using the Joystick ($t(38) = -0.179, p = 0.858$), but Speed Circle showed slightly better times in the Target practice ($t(38) = 2.596, p = 0.013$).
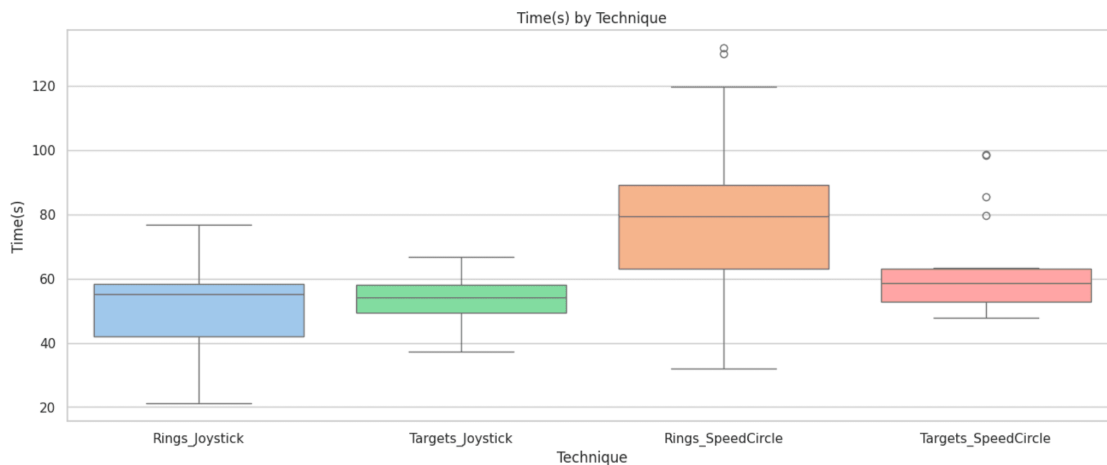
Figure 8. Time to complete each task (Ring and Target) with each speed technique (Joystick and Speed Circle).

Figure 8 shows that for the Target+Joystick task, the median completion time is approximately 55 seconds, with a mean slightly above the median, indicating a right-skewed distribution. The IQR is relatively narrow, suggesting consistent performance among participants. In the Target+Speed Circle task, the median completion time is around 60 seconds, with a mean close to the median, indicating a relatively symmetric distribution and consistent performance, with a few outliers. The Rings+Joystick task has a median completion time of about 50 seconds, with a mean close to the median. The IQR is narrow and shows consistent performance, although a few outliers indicate some variability. Rings+Speed Circle has the highest median completion time at around 80 seconds, with a mean slightly higher than the median, indicating a right-skewed distribution. The IQR is the widest among the conditions, suggesting significant variability in performance, with several outliers showing that some participants took much longer to complete the task.

(a) Flying vs. idle   **(b)** Path Length   **(c)** Overall Speed   **(d)** Collisions

Figure 9. Comparison of performance metrics across joystick and speed circle techniques, including ring navigation and target shooting tasks, to evaluate multitasking performance during VR navigation: (a) percentage of time in motion (Flying vs. Idle), (b) total path length (units), (c) average navigation speed (m/s), and (d) total collision time (seconds).

Flying vs. Idling

See Figure 9a. Speed circle tasks had higher flying percentages for both Ring (87.072, $t(38) = -4.921, p < 0.001$) and Target tasks (87.928, $t(38) = -6.259, p < 0.001$) than the joystick method. Idle time was significantly higher with the joystick, indicating more consistent movement with the speed circle. Adding multitasking did not affect movement fluidity, as shown by no significant difference in flying and idle percentages between tasks using the joystick ($t(38) = -1.073, p = 0.290$) or speed circle ($t(38) = -0.268, p = 0.790$).

Path Length

See Figure 9b. Path length averages were similar among tasks and Ring conditions, ranging from 376.67 to 408.86 meters. There are no significant differences in Ring tasks between joystick and speed circle techniques, with users traveling approximately 2.37% less using the joystick ($t(38) = -0.676, p = 0.503$), consistent with the Magic Carpet study (see further below). No significant difference in path lengths was found in the target practice task ($t(38) = 1.303, p = 0.200$).

Speed

The speed plot (Figure 9c) shows that Rings+Joystick had a median speed of $3.5m/s$ with high variability. Rings+Speed Circle shows a lower median speed of $2m/s$ with slightly more consistency, probably because users did not always reach the maximum speed. Targets+Joystick has a median speed of $2.5m/s$ with high variability, while Targets+Speed Circle has the highest median speed at $3.5m/s$, also with high variability. The joystick technique supports higher speeds, especially in Rings tasks.

Collisions

The collisions did not show significant differences (See Figure 9(d)) between the techniques in the rings ($t(38) = 0.369, p = 0.714$) and the target practice ($t(38) = 0.186, p = 0.853$) tasks. More collisions occurred in the Target practice task than in the Rings task, indicating a slight decrease in performance with multitasking. This pattern holds for both joystick ($t(38) = -2.067, p = 0.045$) and speed circle ($t(38) = -2.569, p = 0.014$) techniques.

5.4. Performance compared to Magic Carpet

The joystick and speed circle techniques covered less distance than the Magic Carpet study (see Figure 10). This occurs even for the target task where a secondary subtask is active. Threading a shorter path indicates that the maneuverability of the technique is more efficient. Table 3 shows the significant differences after applying a Kruskal–Wallis H test and a Dunn test adjusted with Bonferroni.
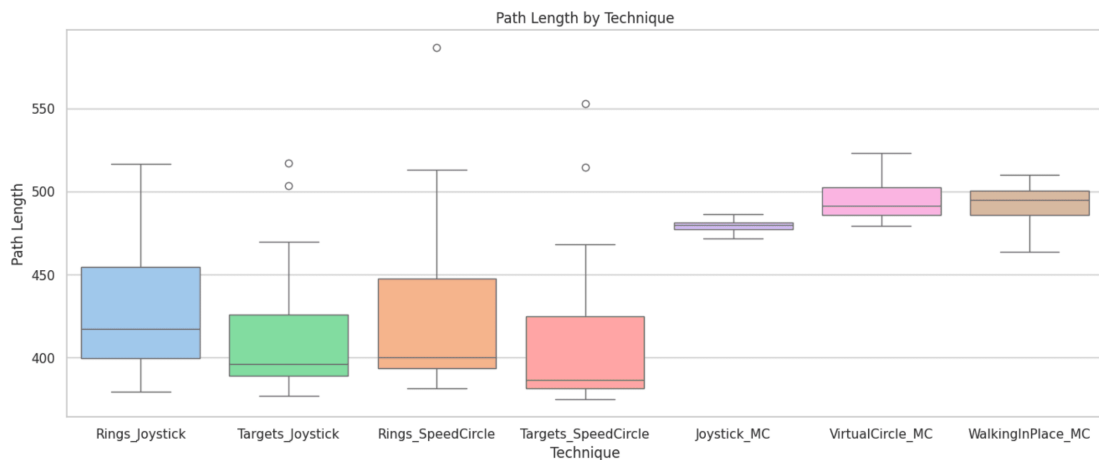
Figure 10. Comparison of the traveled path with each of our conditions and the Magic Carpet three conditions.

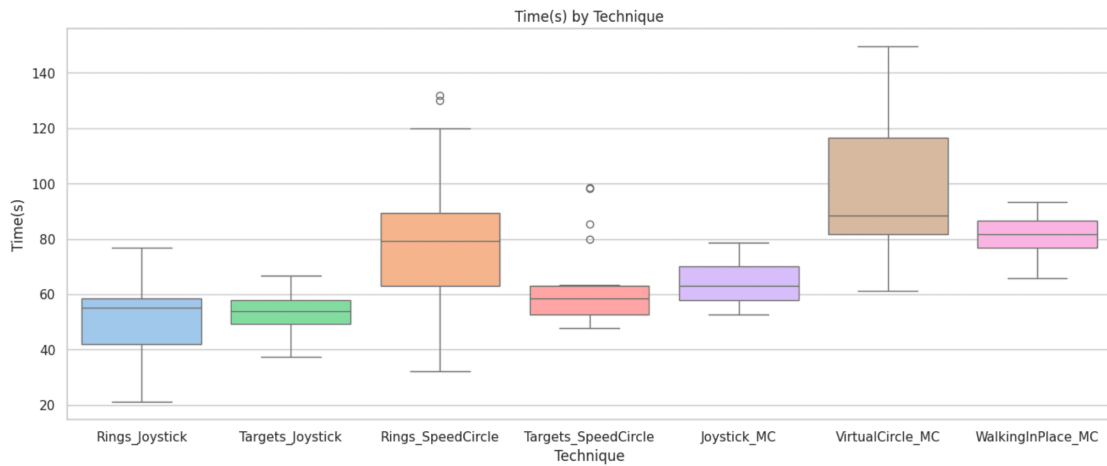|	Rings Joy	Targets Joy	Rings SC	Targets SC	Joy MC	VC MC	WIP MC
Rings Joy	1.0000	1.0000	1.0000	1.0000	0.7339	**0.0017**	**0.0052**
Targets Joy	1.0000	1.0000	1.0000	1.0000	**0.0300**	**0.0000**	**0.0000**
Rings SC	1.0000	1.0000	1.0000	1.0000	0.3482	**0.0004**	**0.0015**
Targets SC	1.0000	1.0000	1.0000	1.0000	**0.0072**	**0.0000**	**0.0000**
Joystick MC	0.7339	**0.0300**	0.3482	**0.0072**	1.0000	1.0000	1.0000
VC MC	**0.0017**	**0.0000**	**0.0004**	**0.0000**	1.0000	1.0000	1.0000
WIP MC	**0.0052**	**0.0000**	**0.0015**	**0.0000**	1.0000	1.0000	1.0000

Table 3. Significance for the Path Length differences among conditions

Although very few collisions occurred overall, there were significantly fewer collisions in the ring task with our technique, either using a joystick or speed circle for speed, than with the Magic Carpet conditions, except for Joystick, where the effect is seen but is insignificant. It is also visible that the target task distracts some participants enough so that they allow more collisions. Still, the distribution is more widespread, showing that some participants are unaffected, implying that training can improve their performance. Significance measures after applying a Kruskal–Wallis H test and a Dunn's Test adjusted with Bonferroni are in table 4.

	Rings Joy	Targets Joy	Rings SC	Targets SC	Joy MC	VC MC	WIP MC
Rings Joy	1.0000	0.2170	1.0000	0.1254	1.0000	**0.0101**	**0.0056**
Targets Joy	0.2170	1.0000	0.1304	1.0000	1.0000	1.0000	1.0000
Rings SC	1.0000	0.1304	1.0000	0.0731	1.0000	**0.0054**	**0.0029**
Targets SC	0.1254	1.0000	0.0731	1.0000	1.0000	1.0000	1.0000
Joy MC	1.0000	1.0000	1.0000	1.0000	1.0000	0.6829	0.4678
VC MC	**0.0101**	1.0000	**0.0054**	1.0000	0.6829	1.0000	1.0000
WIP MC	**0.0056**	1.0000	**0.0029**	1.0000	0.4678	1.0000	1.0000

Table 4. Significance for the differences of collision events among conditions

Finally, the results of the time performance show that our eye gaze steering technique, either with a joystick or a speed circle, is more efficient than the Magic Carpet conditions or at least similar, but not worse (cf. Fig 11). Although $Joystick\_MC$ is not significantly different from our joystick results (see Table 5 for significance), this is due to the lack of training with our technique, which causes the samples to spread. However, all participants in our experiment performed the rings before the target task. This induced a learning effect that considerably reduced variance. This also reduced the average time for the speed circle, which requires more learning than the joystick.

Figure 11. Comparison of the time to complete the task with each of our conditions and the Magic Carpet's three conditions.

	Rings Joy	**Targets Joy**	**Rings SC**	**Targets SC**	**Joy MC**	**VC MC**	**WIP MC**
Rings Joy	1.0000	1.0000	**0.0018**	1.0000	0.4893	**0.0000**	**0.0000**
Targets Joy	1.0000	1.0000	**0.0010**	1.0000	0.3412	**0.0000**	**0.0000**
Rings SC	**0.0018**	**0.0010**	1.0000	0.5095	1.0000	0.7758	1.0000
Targets SC	1.0000	1.0000	0.5095	1.0000	1.0000	**0.0004**	**0.0256**
Joy MC	0.4893	0.3412	1.0000	1.0000	1.0000	**0.0114**	0.2470
VC MC	**0.0000**	**0.0000**	0.7758	**0.0004**	**0.0114**	1.0000	1.0000
WIP MC	**0.0000**	**0.0000**	1.0000	**0.0256**	0.2470	1.0000	1.0000

Table 5. Significance measures for the time to complete differences among conditions

5.5. Discussion

The usability testing of our VR steering system yielded highly promising results, as evidenced by the metrics from the SUS–presence, SSQ, and NASA-TLX questionnaires. The SUS scores indicated a strong presence, with an average score of 15.54 out of 20 (77.7%). This suggests that our system provides a high level of user immersion. The SSQ score, with an average total score of 21.25 could be significant, especially when disorientation was scored higher, at 30.37. However, the nausea factor is lower, 13.87 on average. These

symptoms are concerning but still lower than similar approaches like the Magic Carpet (with joystick SSQ= 39.52; speed circle SSQ=43.84). This indicates that overall comfort was maintained.

The NASA-TLX total score of 28.91, places the system within the medium workload category. Since multitasking was expected to overload the participant's work further, this result indicates that gaze-hand steering successfully accommodated multitasking.

There were significant differences in flying and idle times between techniques, with longer idle times occurring during both tasks when using the joystick compared to the speed circle. This difference was more pronounced in the Target task. The joystick caused users to stop moving more frequently to look around. This can be attributed to the ease of performing dual tasks with body gestures rather than physical input.

There are indications of slightly lower performance in reaching the rings using the speed circle. Regarding the Target task, the mean number of targets destroyed using the speed circle (28.95) is very similar to the joystick (28.85). This indicates that the technique does not affect the accuracy of the Target task.

Results compared with the Magic Carpet technique show that our technique allows for overall performance similar to the previous technique while allowing for a secondary simultaneous task. Some performance measures are better with our technique, such as collisions and distance traveled, arguably due to the speed circle and HMD setup improvements.

Regarding body gestures, we argue that evaluating the user's position within the circle could be more accurately achieved by assessing the position of the torso rather than the head. Users reported difficulties looking upwards and attempting to navigate forward as their heads instinctively moved back. This resulted in a shorter distance from the circle's center, leading to slower velocities.

6. Conclusion

In this paper, we approached the problem of how a gaze-directed approach assisted by hand pointing could prevail relative to other known steering techniques. We designed a method that provides a free look, opening possibilities for multitasking during travel in virtual environments.

We developed an application to study our technique and conducted a user study with 20 participants. Additionally, we analyzed our gaze-directed approach compared to the Magic Carpet to find any possible effects. The results help us understand how applying this method in eye-gaze-dominant applications affects exploration by decoupling gaze from directed steering while still being able to navigate simultaneously.

One important finding is that performance does not decrease when introducing a new task while navigating. Users were able to complete objectives like passing through rings and shooting targets with success rates of

97.5% and 96.3%, respectively, which confirms the effectiveness of our gaze-directed technique. Based on these results, we argue that the system can be used to navigate virtual environments in a range of applications.

One limitation of our study is the collection of user cybersickness, presence, and workload inputs for each technique. The current results assess these variables in general rather than for specific techniques or tasks. A more detailed analysis could improve comparisons with the Magic Carpet and other studies in future work to provide clearer conclusions. An important feature to implement is to use the torso instead of the head position for speed control. We also suggest that hand-tracking instead of controllers be used in a new user study, as this would align with the current trend of gaze-dominant hands-free interaction in XR.

Statements and Declarations

Privacy/Ethics Statement

This study, approved by an ethics board, examines a gaze-directed VR navigation technique for multitasking, using anonymized eye tracking data. Although our method protects privacy, broader applications should consider data sovereignty and informed consent to mitigate unintended profiling or misuse risks.

Footnotes

1 https://www.apple.com/apple-vision-pro/

References

1. a, b, c, d, e, f, g, h, i*Medeiros D, Sousa M, Raposo A, Jorge J (2020). "Magic Carpet: Interaction Fidelity for Flying in V R". IEEE Transactions on Visualization and Computer Graphics. 26 (9): 2793-2804. doi:10.1109/TVCG.2019.2905200.*

2. ^*LaViola Jr JJ, Kruijff E, McMahan RP, Bowman D, Poupyrev IP. 3D user interfaces: theory and practice. Addison-Wesley Professional; 2017.*

3. a, b, c*Jacob RJ. What you look at is what you get: eye movement-based interaction techniques. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '90. New York, NY, USA: Association for Computing Machinery; 1990. p. 11–18. doi:10.1145/97243.97246.*

4. a, b, c, d, e*Kang S, Jeong J, Lee GA, Kim SH, Yang HJ, Kim S. The RayHand Navigation: A Virtual Navigation Method with Relative Position between Hand and Gaze-Ray. In: Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems; 2024; Honolulu, HI, USA. New York, NY, USA: Association for Computing Machinery. doi:10.1145/3613904.3642147.*

5. a, b, c, d *Chen DL, Giordano M, Benko H, Grossman T, Santosa S. "GazeRayCursor: Facilitating Virtual Reality Target Selection by Blending Gaze and Controller Raycasting." In: Proceedings of the 29th ACM Symposium on Virtual Reality Software and Technology; 2023; Christchurch, New Zealand. New York, NY, USA: Association for Computing Machinery. VRST '23. doi:10.1145/3611659.3615693.*

6. a, b *Wagner U, Lystbe6k MN, Manakhov P, Grf8nbe6k JES, Pfeuffer K, Gellersen H. A Fitts2019 Law Study of Gaze-H and Alignment for Selection in 3D User Interfaces. In: Proceedings of the 2023 CHI Conference on Human Factors i n Computing Systems. New York, NY, USA: Association for Computing Machinery; 2023. CHI '23. doi:10.1145/35445 48.3581423.*

7. a, b, c *Lystb\ae{}k MN, Pfeuffer K, Gr\o{}nb\ae{}k JES, Gellersen H (2022). "Exploring Gaze for Assisting Freehand S election-based Text Entry in AR". Proc. ACM Hum.-Comput. Interact.. 6: 141. doi:10.1145/3530882.*

8. a, b *Lystbe6k MN, Rosenberg P, Pfeuffer K, Grf8nbe6k JE, Gellersen H (2022). "Gaze-Hand Alignment: Combining Ey e Gaze and Mid-Air Pointing for Interacting with Menus in Augmented Reality". Proc. ACM Hum.-Comput. Interac t.. 6: 145. doi:10.1145/3530886.*

9. a, b *Suma EA, Babu S, Hodges LF. "Comparison of travel techniques in a complex, multi-level 3d environment." In: P roceedings of the 2007 IEEE Symposium on 3D User Interfaces. IEEE; 2007.*

10. ^ *Suma EA, Finkelstein SL, Reid M, Ulinski A, Hodges LF. "Real walking increases simulator sickness in navigationa lly complex virtual environments." In: Proceedings of the 2009 IEEE virtual reality conference. IEEE; 2009. p. 245–246.*

11. ^ *McMahan RP, Bowman DA, Zielinski DJ, Brady RB (2012). "Evaluating Display Fidelity and Interaction Fidelity in a Virtual Reality Game". IEEE Transactions on Visualization and Computer Graphics. 18 (4): 626–633. doi:10.1109/ TVCG.2012.43.*

12. ^ *Dam P, Braz P, Raposo A. A Study of Navigation and Selection Techniques in Virtual Environments Using Microso ft Kinect®. In: Shumaker R, editor. Virtual Augmented and Mixed Reality. Designing and Developing Augmented a nd Virtual Environments. Berlin, Heidelberg: Springer Berlin Heidelberg; 2013. p. 139-148. ISBN 978-3-642-39405-8.*

13. ^ *Bruno L, Sousa M, Ferreira A, Pereira JM, Jorge J (2017). "Hip-directed walking-in-place using a single depth came ra". International Journal of Human-Computer Studies. 105: 1–11. doi:10.1016/j.ijhcs.2017.03.006.*

14. ^ *Lai C, Hu X, Segismundo A, Phadke A, McMahan RP. "The Comfort Benefits of Gaze-Directed Steering." In: Procee dings of the 2020 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct), 20 20. pp. 102-103. doi:10.1109/ISMAR-Adjunct51615.2020.00040.*

15. ^ *Suma E, Finkelstein S, Reid M, Babu S, Ulinski A, Hodges LF (2009). "Evaluation of the cognitive effects of travel t echnique in complex real and virtual environments". IEEE Transactions on Visualization and Computer Graphics. 16 (4): 690–702.*

16. ^Bowman DA, Koller D, Hodges LF. "Travel in immersive virtual environments: an evaluation of viewpoint motion control techniques." In: Proceedings of IEEE 1997 Annual International Symposium on Virtual Reality; 1997. p. 45–5 2. doi:10.1109/VRAIS.1997.583043.

17. ^Christou C, Tzanavari A, Herakleous K, Poullis C. "Navigation in virtual reality: Comparison of gaze-directed and pointing motion control." In: Proceedings of the 2016 18th Mediterranean Electrotechnical Conference (MELECO N). IEEE; 2016. p. 1-6.

18. a, b Zeleznik RC, Forsberg AS, Schulze JP (2005). "Look-that-there: Exploiting gaze in virtual reality interactions". B rown Univ., Providence, RI, USA, Tech. Rep. CS-05.

19. ^Lai C, Hu X, Aiyaz AA, Segismundo A, Phadke A, McMahan RP (2021). "The Cognitive Loads and Usability of Tar get-based and Steering-based Travel Techniques". IEEE Transactions on Visualization and Computer Graphics. **27** (11): 4289-4299. doi:10.1109/TVCG.2021.3106507.

20. ^Habgood MPJ, Wilson D, Moore D, Alapont S (2017). "HCI Lessons From PlayStation VR." In: Extended Abstracts P ublication of the Annual Symposium on Computer-Human Interaction in Play, CHI PLAY '17 Extended Abstracts, A msterdam, The Netherlands, p. 125–135. doi:10.1145/3130859.3131437.

21. ^Zielasko D, Law YC, Weyers B. "Take a Look Around ⊠ The Impact of Decoupling Gaze and Travel-direction in Se ated and Ground-based Virtual Reality Utilizing Torso-directed Steering." In: Proceedings of the 2020 IEEE Confer ence on Virtual Reality and 3D User Interfaces (VR); 2020. p. 398-406. doi:10.1109/VR46266.2020.00060.

22. a, b Tregillus S, Al Zayer M, Folmer E. Handsfree Omnidirectional VR Navigation Using Head Tilt. In: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems; 2017; Denver, Colorado, USA. New York, NY, US A: Association for Computing Machinery; 2017. p. 4063–4068. doi:10.1145/3025453.3025521.

23. ^Medeiros D, dos Anjos RK, Mendes D, Pereira J\u00e3o M, Raposo A, Jorge J. Keep my head on my shoulders! why third-person is bad for navigation in VR. In: Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology; 2018; Tokyo, Japan. New York, NY, USA: Association for Computing Machinery. doi:10.1145/328150 5.3281511.

24. ^Slater M, Usoh M, Steed A (1994). "Depth of Presence in Virtual Environments". Presence: Teleoperators and Virtu al Environments. **3** (2): 130–144. doi:10.1162/pres.1994.3.2.130. PDF.

25. ^Kennedy RS, Lane NE, Berbaum KS, Lilienthal MG (1993). "Simulator Sickness Questionnaire: An enhanced meth od for quantifying simulator sickness". The International Journal of Aviation Psychology. **3** (3): 203–220.

26. ^Hancock PA, Meshkati N. Human mental workload. Vol. 52. North-Holland Amsterdam; 1988.

Declarations

Potential competing interests: No potential competing interests to declare.