Eyes-free Target Acquisition During Walking in Immersive Mixed Reality

Qiushi Zhou, Difeng Yu, Martin N Reinoso, Joshua Newn, Jorge Goncalves, Eduardo Velloso

Abstract—Reaching towards out-of-sight objects during walking is a common task in daily life, however the same task can be challenging when wearing immersive Head-Mounted Displays (HMD). In this paper, we investigate the effects of spatial reference frame, walking path curvature, and target placement relative to the body on user performance of manually acquiring out-of-sight targets located around their bodies, as they walk in a spatial-mapping Mixed Reality (MR) environment wearing an immersive HMD. We found that walking and increased path curvature negatively affected the overall spatial accuracy of the performance, and that the performance benefited more from using the torso as the reference frame than the head. We also found that targets placed at maximum reaching distance yielded less error in angular rotation and depth of the reaching arm. We discuss our findings with regard to human walking kinesthetics and the sensory integration in the peripersonal space during locomotion in immersive MR. We provide design guidelines for future immersive MR experience featuring spatial mapping and full-body motion tracking to provide better embodied experience.

Index Terms—Mixed Reality, Virtual Reality, Target Acquisition, Motion Tracking, Proprioception, Locomotion, Sensory Integration

◆

1 INTRODUCTION

In daily life, we can easily reach and grab objects near our bodies during locomotion even without looking at them—for example, pulling a ticket out of the rear pocket while boarding a bus or reaching for the phone to change a song while jogging. These tasks demand the unconscious integration of vision and proprioception to infer our movement velocity in relation to the external environment and the awareness of the locations of different body parts relative to each other. However, the ability to perform these tasks is hindered in immersive mixed reality (MR): our vision is replaced by the rendering of the virtual environment (VE) through the head-mounted display (HMD), while proprioception from the moving limbs remains unaltered. This challenge is further increased by the fact that in a VE, users are no longer restricted to placing targets *on* the body, but rather, they are free to place them anywhere *around* the body and anchor them to any body part.

As a motivating example, consider a hypothetical blueprint-editing application that uses the entire room as the canvas for controller-free interaction with motion tracking (Figure 1). The engineers are able to grab different tools surrounding them while walking to the locations of interest, as they rely on vision of the VE to know their own locations and rely on proprioception to direct their hands towards the intended tools, which are anchored to their bodies in the space out of the limited field-of-view (FoV). Compared with menu-based interaction, this arrangement enables faster acquisition relying on proprioception while users direct their visual attention on the task. Additionally, it saves the extra head movement for finding targets around the body to overcome the limited FoV, minimising motion sickness [26]. Being able to accurately locate body parts during locomotion in immersive VE helps with any field of application featuring body movement tracking. For example, players of first-person shooting games can quickly trigger weapons equipped around their avatars in the VE while fixating their gaze on the enemy and moving around to dodge attacks. Similarly, choreographers would be able to design experimental performances in immersive VE while relying on the dancers' bodily perception for their movement quality. In this work, we identify three design factors and investigate their effect on the performance of eyes-free target acquisition.

Fig. 1. Motivating use case: A room-scale blueprint-editing interface enables engineers to grab and use the tools surrounding them during walking. The engineer in the figure grabbed the measurement tool with the right hand as he looked at an area at his upper-left side, while his right hand being occupied by holding physical document. This interface allows the engineer to always keep their attention focused on the task.

During eyes-free target acquisition tasks, our motor control mechanism responds to stimuli within the Peripersonal Space (PPS)—a multi-sensory representation of the reachable space around different parts of the body [46]. Previous works have found that we adaptively use different parts of the body, such as the face and the torso, as a reference frame for the PPS according to the stimuli [47]. Because the perception of vision and proprioception are separated when wearing an immersive HMD, it is unclear whether accurate target acquisition benefits more from using the head or the torso as the **reference frame** for perceiving the spatial layout of the targets within the PPS, notably when their orientations can differ up to 25° when walking along a curved path [10]. Apart from increased head rotation, body movements associated with different **locomotion patterns**, such as lateral oscillations of the body due to the gait rhythm and restricted arm swings during reaching actions, can further hinder performance. Finally, though re-

---

- *The authors are with School of Computing and Information Systems, The University of Melbourne. E-mails: {qiushiz2, difengy, mreinoso, jnewn} @student.unimelb.edu.au; {jorge.goncalves, eduardo.velloso} @unimelb.edu.au.*

Manuscript received 18 May 2020; revised 26 July 2020; accepted 17 Aug. 2020. Date of publication 17 Sept. 2020; date of current version 3 Nov. 2020. Digital Object Identifier no. 10.1109/TVCG.2020.3023570

search on PPS has found that people are able to perceive the locations of their hands more accurately when they are closer to the body [22], other evidence suggests that the increased rotation of the joint angles makes it more challenging to accurately locate near-body objects than using the stretched-out arm to acquire objects at reaching distance while relying on proprioception [32]. Thus, **target placement** in near- versus far-PPS is another potential factor affecting the performance of eyes-free target acquisition.

In this paper, we investigate the effect of **reference frame** of the PPS, **locomotion patterns** that update the PPS, and **target placement** relative to the boundary of the PPS on user performance for manually acquiring out-of-sight targets located in their PPS, as they walk in a spatial-mapping MR environment wearing an immersive HMD. We found that locomotion negatively affected the spatial accuracy of the task, especially when the curvature of the walking path increased. We also found that the overall user performance benefited more from using the torso as the reference frame than from using the head, while target placement at maximum reaching distance yielded better angular- and depth-accuracy of the acquisition. We contribute an understanding of the sensory integration mechanism behind eyes-free target acquisition during walking in immersive MR, and provide design guidelines for future immersive MR experience featuring spatial mapping and full-body motion tracking.

2 RELATED WORK

2.1 Eyes-Free Spatial Interaction in MR

The reachable space around the body—the peripersonal space (or PPS)—is an important region for mixed reality (MR) and other mid-air gestural interfaces. Most modern HMDs, such as the HoloLens[1] and Magic Leap[2] include some form of motion tracking and/or gesture recognition, enabling manual interaction in the PPS. Previous work has envisioned using the PPS as a ‘wearable information space’ in MR where virtual content is rendered through an HMD following users’ body movement—as if it is ‘worn’ by them [4]. For instance, Ens et al. proposed the *Personal Cockpit*, a spatial MR interface where users manually select virtual windows fixed around their bodies and rendered through the HMD. Users were able to reduce 40% of the time spent for task-switching using this pointing technique than using traditional gestural methods [16]. Thanks to their proprioception and spatial memory, users are able to achieve many tasks that rely on the accurate perception of the locations of their hands in space such as dancing and juggling, with or without looking [9]. This ability shows the potential of accurate target acquisition within the ‘wearable’ PPS in an eyes-free manner [34].

As an intuitive input technique, eyes-free interaction has been studied in various types of applications. One instance is the case of *Virtual Shelves*, which allowed users to open applications on their smartphones by pointing them towards specific directions relative to their bodies without visual aid. Their evaluation showed that users were able to learn and remember the angular locations in space after practice [31]. In another study, Yan et al. found that eyes-free target acquisition in Virtual Reality (VR) was faster than the eyes-engaged method and provided satisfying accuracy in distinguishing angular differences between targets with less fatigue and sickness [55]. Cockburn et al. proposed *Air Pointing*, an interaction technique potentially supporting spatial target acquisition in MR by pointing a controlling device in different virtual cubical grids allocated in space. They found that the performance of acquisition was better when there was only one layer of objects around the body than when multiple targets were aligned along the depth axis [9].

Whereas these works similarly found that users are able to distinguish angular differences between invisible around-body objects with training, they also suggest that users’ depth perception may not be as accurate for the same task. Lubos et al. presented and evaluated a spatial interaction method for VR that leveraged *joint-centred kinespheres*. A kinesphere is the kinematics equivalent to the reachable PPS around joints, such as shoulder, elbow, and wrist. They found that targets placed on the boundary of the kinespheres of the joints yielded better performance than at locations closer to the body [32]. However, previous works have also found that spatial perception is generally more accurate when targets are closer to the body in the PPS and that more extreme joint positions are overestimated [42, 53]. It is thus unclear if the benefits of placing targets at maximum reaching distance with minimum bending of the arm would outweigh the benefit of placing targets closer to the body in more comfortable and accessible positions of users’ own choices. In this work, we investigate this effect of target placement with varying distance relative to the body on eyes-free target acquisition performance in the context of walking in immersive MR.

2.2 Target Acquisition During Walking in Immersive MR

As real walking has not been universally available in VR or immersive MR until recently, most research on locomotion in those environments has been based on substitutive walking interfaces devised to simulate real walking, including traditional or omnidirectional treadmills [11, 12, 48], stepping or leaning motions [27], and redirected walking [40]. Whereas the embodied sensation of self-motion is naturally experienced in real walking, none of the substitutive technologies provide a sensory experience as compelling [14, 49]. This is because a realistic walking experience involves body movements other than direct translation, such as rhythmic oscillations and lateral movements of the body naturally induced by the whole-body stepping motion and arm swings [10, 29, 33].

Though contributing to realistic embodied experience, these body movements can also hinder the concurrent performance of tasks that demand accuracy of body movement in the space. For instance, we unconsciously swing our arms during walking to compensate forward stepping motions of the lower limbs. Previous works on walking dynamics have found that to maintain gait stability, deliberate modification of arm-swing amplitude affects the walking velocity and the energy expenditure of the walking movement [15, 33, 51]. When people reach toward targets around them during walking, the movements of the reaching arms are superimposed over natural arm swings [7]. Consequently, the reaching movement changes concurrent movement patterns of the torso and the lower limbs, such as the stride length, the walking speed, and the rhythmic oscillation of the body [7, 8]. Additionally, this modulation of whole-body movement can be further complicated by concurrent walking and turning, when the rhythmic body oscillations are shifted inward in relation to the walking path, causing increased angular acceleration and decreased velocity of the body [10]. In this work, we compare user performance of manual target acquisition during standing-still and walking with different path curvatures.

Walking-induced body movements have specific effects on how we update our spatial perception during turning. Previous works have found that in order to maintain gaze stability during walking and turning, the head turns towards the inner part of the walking path more than the torso does for as much as 25° [10, 29]. The effect of this discrepancy between head- and torso-orientation during walking and turning is amplified in immersive MR environments because it adds to the difference between how head- and torso- movements are perceived. When wearing an immersive HMD, the wearer predominantly perceives their own head movements by seeing the corresponding changes in their view of the virtual environment through the HMD, while they perceive the movements of the torso and the limbs through their proprioception [3]. Previous works have found that the PPS can be centred around different body parts such as the torso and the face, and that the body-part-centred coordinates provide the general solution to sensory-motor integration that guide users in locating the sensory stimuli around them [22, 28]. To sum up, the literature suggests that using different frames of reference, i.e. head-centred or body-centred, may have different impact on the accuracy of locating targets in the PPS in the context of immersive MR. We identify this inherent problem and compare the accuracy of eyes-free target acquisition between using the head and using the torso as reference frame of the targets around the user’s body during walking.

2.3 Sensory Integration in the Peripersonal Space

The term “peripersonal space” denotes a mental representation of the space around users’ bodies that guide their motor actions as a motor-

---

[1] https://docs.microsoft.com/en-us/hololens/hololens1-basic-usage

[2] https://developer.magicleap.com/en-us/learn/guides/design-gesture

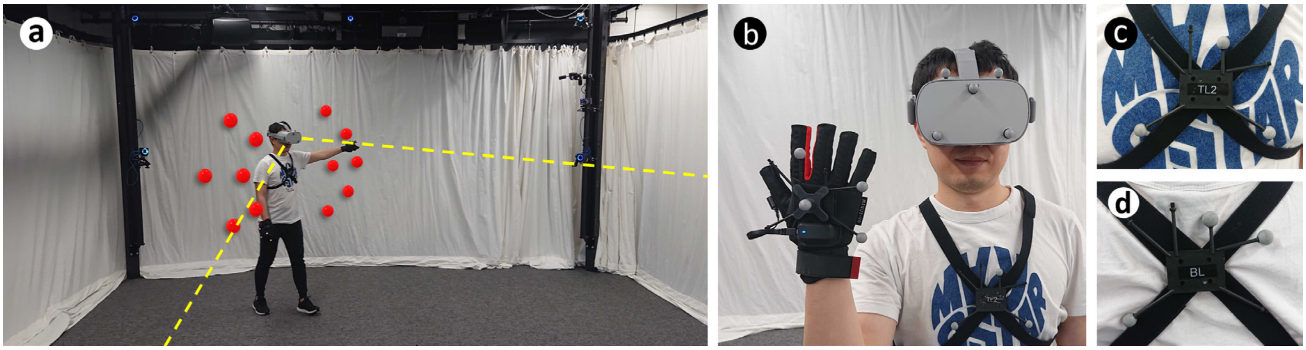


Fig. 2. Study Setup: (a) participant performing target acquisition (grasping gesture) while walking in the tracked space; (b) reflective position-tracking markers attached on the immersive HMD and on the gloves, and (c,d) on the front- and back-side of the participant's torso.

to-sensory pathway for the construction of object and space perception within reachable distance [28,43,44]. During walking, the PPS gets constantly updated following the movements of different parts of the body as a real-time map to guide action toward reachable objects [22, 24]. Because the construction of the PPS relies on multi-sensory information, we must understand how movement-induced change in sensory information from different modalities are integrated by the brain before we can understand how and why the movements of different parts of the body involved in the walking and reaching behaviour may affect target acquisition performance [6].

We rely on vision and proprioception for perceiving both egocentric and allocentric spatial information. Vision provides us with movement information relative to the environment such as the direction of the movement and of the head, in the form of optic flow—the temporal change in vision following head- and eye-movement [52]. Proprioception provides us with the sense of locations of and relative movements between different parts of our bodies where sensory receptors are located [50]. Previous works have found that human perception of self-motion relies on the integration of visual and proprioceptive feedback [20, 35, 36]. It has also been found that arm positions are represented in the primate brain by integrating visual and proprioceptive information onto the same neurons which respond to the felt positions of the arms when they are out-of-sight [21].

When vision and proprioception are in conflict while perceiving movement, they are integrated following a weighted model depending on the reliability of each modality in the specific context [17–19]. Previous works have found that proprioception dominates visual information in those conflicting situations, and that optic flow by itself almost always have worse performance than proprioception in terms of perceiving walking/turning speed, distance travelled, body orientation, and spatial interaction in real and immersive virtual environments [1, 18, 25, 30, 39]. In immersive MR, the processing of visual and proprioceptive sensory information are separated due to the HMD worn by users. To better understand the mechanisms behind how eyes-free target acquisition performance is affected by the factors of interest, we discuss our results with regards to previous works on sensory integration within the PPS.

In this work, we investigate the effects of target placement, reference frame, and walking path curvature on the performance of eyes-free target acquisition in immersive MR. We compare the performance between the condition where participants grasped targets located at their maximum reaching distance and the condition where participants were allowed to move the targets closer to their bodies in more comfortable positions within a small range. We further compare the performance between the contexts of standing-still, rectilinear walking (i.e. straight line) and curvilinear walking (8-shaped path). We also compare the performance between using the head and using the torso as reference frame for placing the targets around the user's body. With the discussion of our results, we offer insights as to how sensory integration during walking in immersive MR affects our spatial perception of the interactive space around us.

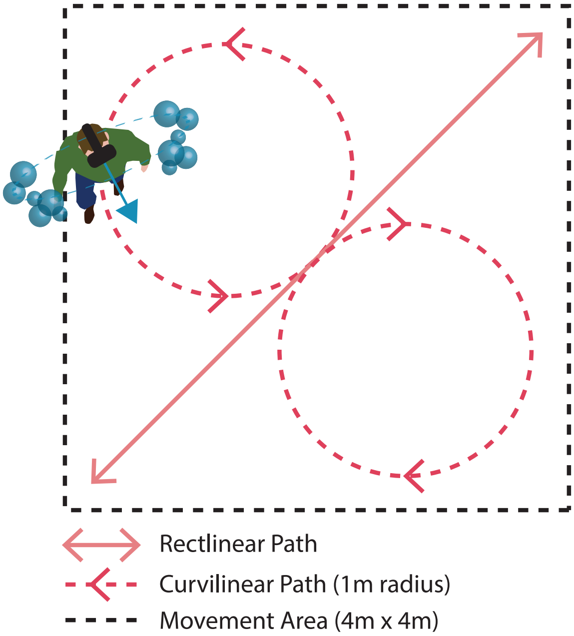


Fig. 3. Diagram showing the walking paths in the tracked lab space. Participants walked following a cue which moved diagonally back and forth between two corners of the movement area during RECTILINEAR walking, and in circular motion during CURVILINEAR walking.

3 METHOD

We designed and conducted a study to investigate the effect of reference frames, locomotion patterns, and target placements on eyes-free target acquisition during walking in a spatial-mapping immersive MR environment. The study was conducted under the approval of the University's Human Ethics Committee.

3.1 Study Design

Previous work has evaluated the performance of eyes-free target acquisition in immersive VE with participants sitting and standing still [55]. Performing the same task during walking in immersive MR environment adds to the task complexity the movements naturally induced by walking, and the separated perceptions of head and body movement as a result of walking motions of the torso and the lower limbs [7, 8, 15, 33, 51]. Additionally, the difference between the spatial orientations of the head and the torso increases with the curvature of the walking path [10, 29]. Consequently, users may adaptively use their head or torso as the spatial reference frame that they perceive the targets to be rotating around [22, 28]. Further, previous works have found that whereas proprioception-based spatial perception within the

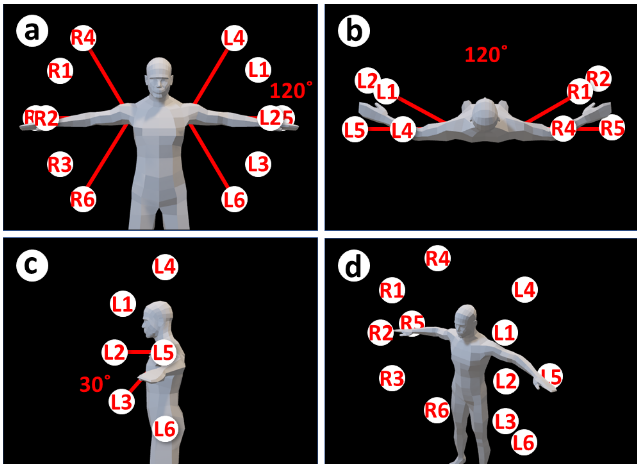


Fig. 4. Illustration of target positions: (a) front view; (b) top view; (c) left view; (d) perspective view with target numbers.

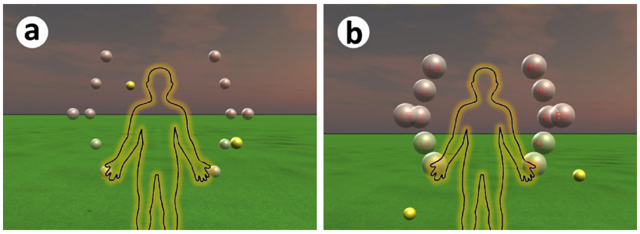


Fig. 5. Targets in the VE: (a) practice round in PRESET and in (b) CUSTOM with visible targets and hands (yellow spheres). The contours of the participant are only displayed here for illustration.

peripersonal space benefit stimuli closer to the body, manually acquiring targets at maximum reaching distance benefits from smaller joint angles and increased neural strength and reliability [32, 38, 42, 53]. However, it is unclear how the addition of walking, the head-torso orientation difference induced by the variation in the path curvature, and the target placement with different distances to the body would affect the performance (Euclidean and angular distances between locations of targets and acquisition points in the 3D space) of eyes-free target acquisition in immersive MR.

We investigated these effects by employing a $3 \times 2 \times 2$ repeated-measures design with three factors: LOCOMOTION, REFERENCE FRAME, and PLACEMENT. We compared performance across three levels of LOCOMOTION patterns: STILL when participants completed the tasks standing still, RECTILINEAR for when participants walked in a straight line, and CURVILINEAR for when participants walked around a curve shaped like the number 8. We also compared the acquisition performance with different REFERENCE FRAMES: HEAD where the around-body targets rotate following the spatial orientation of the head, and TORSO where the targets rotate following the orientation of the torso (Figure 6). We investigated the PLACEMENT with two levels: PRESET where targets are located at participants' maximum reaching distance, and CUSTOM where participants chose the distance to the targets themselves, while the angular locations of the target relative to the body remained the same as in the PRESET condition (Figure 5).

Previous works measured the accuracy of locating targets within the PPS by measuring the Euclidean error of target acquisition and the perceived target distance [28, 38]. To quantify the target acquisition performance, we first measured EUCLIDEAN OFFSET, and then separated it into HORIZONTAL ANGULAR OFFSET, VERTICAL ANGULAR OFFSET, and REACH OFFSET to investigate the individual effects of these components, and to gather an in-depth understanding of how the representation of the PPS is updated as an integral space during walking in different patterns [37].

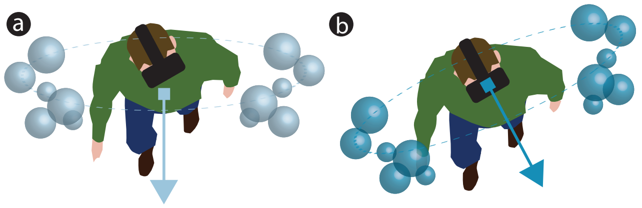


Fig. 6. While the targets are always anchored around the centre of the torso, they rotate differently when using the torso (a) or the head (b) as the reference frame.

3.2 Experimental setup

The study took place within a 4m×4m area in a 6m×6m lab space covered by OptiTrack motion tracking cameras (Figure 2a). Participants wore an Oculus Go[3] wireless HMD and a pair of Manus VR gloves[4]. Participants selected objects with whole-hand grasping gestures tracked by the gloves, resembling how we interact with real-world objects. We attached 19 reflective markers on the front and the top of the HMD, on the front and back of participants' bodies with velcro straps, and on the VR gloves, to track the locations and orientations of their head, torso, and hands independently (Figure 2). In the VE, participants' torso locations were defined by the middle point between the two reflective markers attached to the front and back of their torsos. Their torso orientations were defined as the 3D vector pointing from the back marker to the front marker. Participants' head locations were calibrated at approximately the mid-point between the eyes, and their head orientations were defined as the orientation of the tracked HMD. Hand locations were tracked with the markers attached on the gloves. We measured the distance between the marker and the centre of the palm when a grasping gesture was performed. We calibrated the hand location represented in the OptiTrack system by shifting it from the marker location to the centre of the hand in the distance measured.

We built the VE using Unity 3D (Figure 5). We mapped the 4m×4m tracked area in the lab with a 1:1 relationship to the virtual environment (Figure 3). We masked the environment with a skybox and a ground material without prominent visual landmarks to avoid providing participants with additional spatial references. We instructed participants to follow a target cue (a number floating at the height of the participants' chest) that moved back and forth between the two footprints on the ground in the RECTILINEAR condition, and moved along the brown path printed on the ground in the CURVILINEAR walking condition, following the same cue. We designed the walking path in the CURVILINEAR condition to be a "infinity" shape to balance the number of left and right turns and to trigger a lasting and steady neural stimuli for the brain to perceive a curved path [10]. Participants were instructed to maintain their walking velocity during the walking conditions while they perform the tasks. In each task condition, the cue prompted participants with the number of the target to be acquired. The target cue moved at a constant velocity of 0.7m/s in the walking conditions and remained still at the centre of the room in the STILL condition. The velocity was chosen after trials of different moving speeds of the cue for participants to follow in pilot studies. Participants reported this velocity as the most comfortable without noticeable motion sickness.

We measured participants' shoulder widths and arm lengths, and calculated their shoulder locations relative to the tracker on the torso. The targets were anchored at the distance of participants' shoulder-width on the left and right side of the forward orientation, which was the head or torso orientation when using each as the reference frame (Figure 4). This was to accurately lay out the targets at maximum reaching distance around the shoulder joints, and to minimise the error caused by unintended head or torso rotation induced by arm movement, as supported by previous work [16]. The 12 targets were labelled L1 to L6 (left) and R1 to R6 (right). The radius of the spheres rendered for representing the targets was 0.1m. Placement details of the targets are

---

[3] https://www.oculus.com/go/

[4] https://manus-vr.com/prime-one-gloves/

illustrated in Figure 4. The target locations were chosen to be within reachable PPS, but out of the field of view (FoV) of the Oculus Go and most other popular VR HMDs[5]. Participants were informed that the targets will remain invisible during the tasks.

3.3 Participants

We recruited 24 participants (12F/12M) with a mean age of 24.3 years ($Min = 19, Max = 43, SD = 6.38$) through university mailing lists. Among these, 19 participants (79%) reported prior experience with MR. For each participant, the study lasted approximately one hour on average, and they each received \$10 AUD gift card for their time.

3.4 Procedure

We welcomed participants upon their arrival and introduced the purpose of our study. Participants read and signed a consent form, reported their demographic data, and their previous mixed reality experience. We informed participants that they might experience simulator sickness during the study and that they were free to pause or withdraw from the study at any point in time without any negative consequences. Next, we fitted the Oculus Go and the reflective markers on the participants. After fitting and calibrating the Manus VR gloves through the procedure in their software interface, we measured the participants' shoulder widths and arm lengths to calculate the shoulder locations to centre the targets around. Following, we briefly demonstrated the practice round and the tasks to the participants in the VE to make sure that they fully understood the tasks.

Participants learned the locations of the targets in a practice round before the tasks by grasping them with their hands by making the yellow balls representing their hands overlap with the balls representing the targets (0.05m error allowed). During the practice round, participants were able to see their hands and the targets around their bodies rendered in front of where they stood, to learn the locations of the targets by grasping them with their hands (Figure 5). The colour of the target changed from clear to pink after a participant successfully grasped it, along with a beeping sound confirming the correct acquisition. The locations where participants grasped were remembered by the system as the new locations of the targets to account for the errors (within the 0.05m range) of participants' initial grasps. Following, we showed participants how the targets rotated differently when using the head and torso as a reference frame, and they practised under both conditions. We asked participants to remember the locations of their hands relative to their bodies as they grasped the targets correctly to recall and grasp them again later [13, 34]. Participants learned the target locations by grasping them one by one. In PRESET, the locations of the targets were fixed at participants' approximate arms' lengths. The targets were of the same sizes as the spheres representing the hands (Figure 5c). In CUSTOM, the targets were shown as larger areas (0.2m radius) around the target locations in PRESET in the initial practice round before tasks. Participants were instructed to grasp anywhere within the areas to confirm the new target locations. For example, if a participant grasped anywhere inside target L1, it would shrink to the size of the hand sphere and remain at the new location relative to the participant's body throughout later tasks in CUSTOM.

We started the study once participants were clear on the tasks. We balanced the order of the tasks between the three factors using a Latin Square. Their task was to grasp out-of-sight virtual targets located around participants' bodies that followed them as they moved around the tracked space. Participants repeated the tasks three times for the appearance of each different cue number. The 12 target numbers appeared following a random order. No feedback was provided during the tasks to indicate correct or wrong acquisition, as we aimed to understand users' natural performance instead of improving their performance during the task.

4 RESULTS

We collected 10,368 data points in total (24 participants × 2 PLACEMENT levels × 2 REFERENCE FRAME levels × 3 LOCOMOTION levels × 12 targets × 3 repetitions). We discarded error trials (246 trials, 2.4%) in which participants reacted to the cue with the wrong hand. Due to the nature of our measurement being Euclidean/Angular offset between target position and acquisition position, we were not able to distinguish between large acquisition errors and acquiring the wrong target with the correct hand. To minimise the effect of the trials in which participants made such mistakes, we removed outliers (164 trials, 1.6%) in which the Euclidean distance between the target position and the acquisition position was above three standard deviations from the mean ($mean + 3sd.$). As a result, 410 data points (4.0%) were discarded, leaving 9,958 data points for analysis.

We applied the Aligned Rank Transform (ART) on the data after identifying non-normal distribution of residuals [2, 54]. Next, we performed a factorial ANOVA on each response variable and post hoc pairwise comparisons with Holm–Bonferroni adjustment to analyse the acquisition accuracy. We present measures of response variables and all the statistically significant interaction effects in the following sections. For HORIZONTAL ANGULAR OFFSET, VERTICAL ANGULAR OFFSET and REACH OFFSET, we measured the absolute values as performance error regardless of the directions (i.e., up/down for vertical angular distance and close/far for reach offset). We present the average of the directional offsets as signed values to aid the interpretation of the results. We also measured and compared reaction time across all conditions but did not find any significant result. We present the results with Figure 7-18, in which the error bars indicate 95% confidence interval.

4.1 Euclidean offset

As a measure of the overall accuracy of the target acquisition performance, we define EUCLIEAN OFFSET as the Euclidean distance between the target and the acquisition locations in metres. For the main effect of REFERENCE FRAME ($F_{1,9923.1} = 34.9, p < .001$), the offset in TORSO ($mean = 0.226, sd = 0.11$) was significantly smaller than in HEAD ($mean = 0.238, sd = 0.12$). For the main effect of LOCOMOTION ($F_{2,9923.3} = 88.6, p < .001$), the offset in STILL ($mean = 0.211, sd = 0.10$) was significantly smaller ($p < .001$) than in RECTILINEAR ($mean = 0.240, sd = 0.12$), which is significantly smaller ($p < .05$) than in CURVILINEAR ($mean = 0.247, sd = 0.12$) (Figure 7). The significant interaction effects were PLACEMENT:REFERENCE FRAME ($F_{1,9923.1} = 4.9, p < .05$), PLACEMENT:LOCOMOTION ($F_{2,9923.1} = 6.0, p < .01$) (Figure 8), and PLACEMENT:REFERENCE FRAME:LOCOMOTION ($F_{2,9923.1} = 4.4, p < .05$) (Figure 9). For PLACEMENT:LOCOMOTION, post-hoc pairwise comparison showed that CURVILINEAR was significantly different ($p < .01$) from STILL and RECTILINEAR (Figure 8).

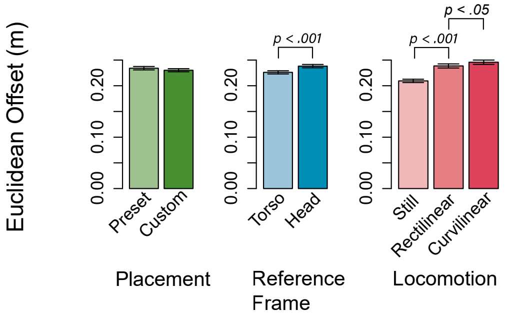


Fig. 7. EUCLIEAN OFFSET between the targets and the acquisition locations in 3D space.

---

[5] 110° was found to be the upper limit of the FoV of most popular HMDs in: https://benchmarks.ul.com/compare/best-vr-headsets

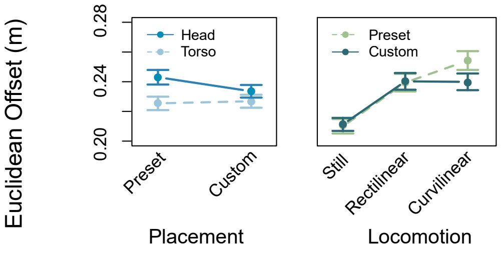
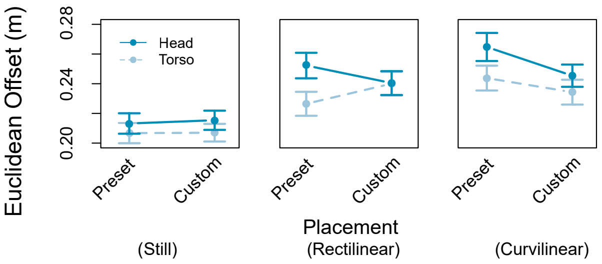


Fig. 8. Two-way interaction effects EUCLIEAN OFFSET between PLACEMENT and REFERENCE FRAME (left), and between PLACEMENT and LOCOMOTION (right).



(Still) (Rectilinear) (Curvilinear)

Fig. 9. Three-way interaction effects in EUCLIEAN OFFSET. The plot is separated in the three LOCOMOTION patterns: STILL (left), RECTILINEAR (middle), and CURVILINEAR (right).

4.2 Reach offset

To measure the offset between the perceived and the actual distance in metres between the targets and the participants' bodies, we define REACH OFFSET as the absolute difference between the Torso-Target Euclidean distance and the Torso-Acquisition Euclidean distance. We found that for PLACEMENT ($F_{1,9924.0} = 144.1, p < .001$), the offset in PRESET ($mean = 0.083, sd = 0.07$) was significantly smaller than in CUSTOM ($mean = 0.102, sd = 0.08$). For LOCOMOTION ($F_{2,9924.1} = 17.5, p < .001$), the offset in STILL ($mean = 0.087, sd = 0.07$) is significantly smaller than in RECTILINEAR ($p < 0.001, mean = 0.095, sd = 0.08$) and in CURVILINEAR ($p < 0.001, mean = 0.095, sd = 0.07$) (Figure 10). We present the signed average of torso-target distance minus torso-acquisition distance to help interpret the results (Figure 11).

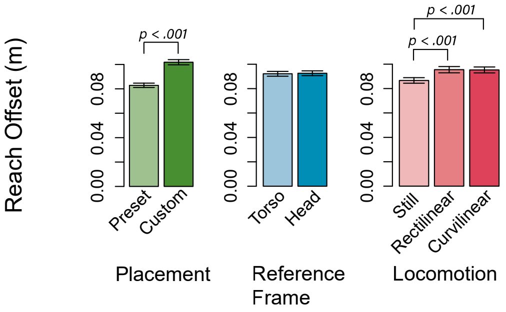


Fig. 10. The absolute difference between the torso-target Euclidean distance and the torso-acquisition Euclidean distance.

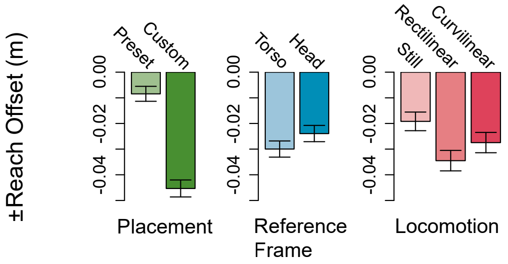


Fig. 11. Average value of torso-target Euclidean distance minus torso-acquisition Euclidean distance.

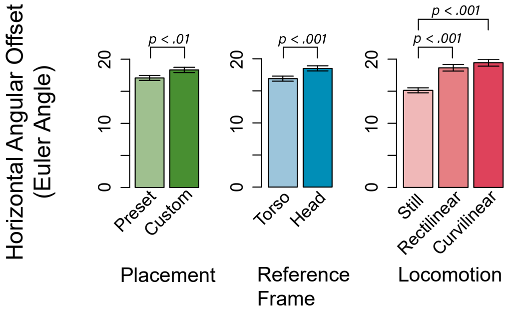


Fig. 12. The absolute horizontal angular difference between the target locations and the acquisition locations around participants' bodies in Euler angle degrees.

4.3 Horizontal Angular offset

We define HORIZONTAL ANGULAR OFFSET as the absolute horizontal Euler angular difference between the target locations and the acquisition locations around participants' torsos. For the main effect of PLACEMENT ($F_{1,9923.2} = 10.5, p < .01$), the offset in PRESET ($mean = 17.087, sd = 13.73$) was significantly smaller than in CUSTOM ($mean = 18.332, sd = 14.82$). For REFERENCE FRAME ($F_{1,9923.1} = 40.6, p < .001$), the offset in TORSO ($mean = 16.902, sd = 14.04$) is significantly smaller than in HEAD ($mean = 18.511, sd = 14.51$). For LOCOMOTION ($F_{2,9923.3} = 64.3, p < .001$), the offset in STILL ($mean = 15.140, sd = 11.70$) is significantly smaller than in RECTILINEAR ($p < 0.001, mean = 18.651, sd = 14.97$) and in CURVILINEAR ($p < 0.001, mean = 19.445, sd = 15.63$) (Figure 12). A two-way interaction effect was found significant in REFERENCE FRAME:LOCOMOTION ($F_{2,9923.2} = 4.8, p < .01$) where post-hoc pairwise comparison showed that REFERENCE FRAME's effect on STILL is significantly different from that on CURVILINEAR ($p < .01$) (Figure 13). We found a significant three-way interaction PLACEMENT:REFERENCE FRAME:LOCOMOTION ($F_{2,9923.1} = 6.2, p < .01$) where post-hoc pairwise comparison showed that RECTILINEAR is significantly different from CURVILINEAR ($p < .01$) and STILL ($p < .01$) in the interaction with the other two factors (Figure 14). We presented the signed average in Figure 15 to aid the interpretation of results.

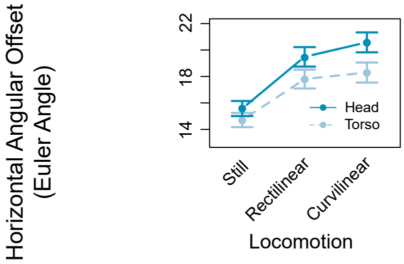


Fig. 13. Two-way interaction effects in HORIZONTAL ANGULAR OFFSET between PLACEMENT and LOCOMOTION.

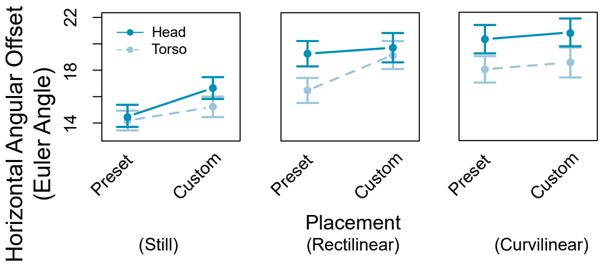


Fig. 14. Three-way interaction effect in HORIZONTAL ANGULAR OFFSET. The plot is separated in the three LOCOMOTION patterns: STILL (left), RECTILINEAR (middle), and CURVILINEAR (right).

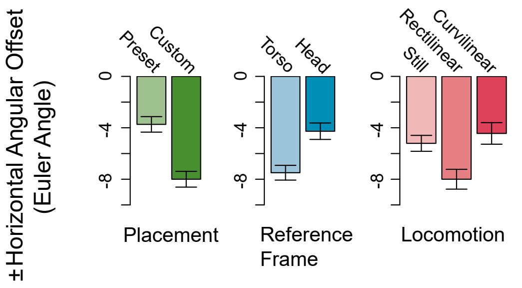


Fig. 15. Signed average of HORIZONTAL ANGULAR OFFSET: the horizontal rotation from the forward vector of the torso to the acquisition location minus the horizontal rotation from the forward vector of the torso to the target location.

4.4 Vertical Angular offset

We define VERTICAL ANGULAR OFFSET as the absolute vertical Euler angular difference between the target locations and the acquisition locations around participants' torsos. For the main effect of PLACEMENT ($F_{1,9923.4} = 3.9, p < .05$), the offset in PRESET (*mean* $= 19.580$, *sd* $= 26.41$) was significantly smaller than in CUSTOM (*mean* $= 19.752$, *sd* $= 24.53$). For the main effect of LOCOMOTION ($F_{2,9923.5} = 16.1, p < .001$), the offset in STILL (*mean* $= 17.181$, *sd* $= 20.87$) was significantly smaller ($p < .05$) than in CURVILINEAR (*mean* $= 19.941$, *sd* $= 26.16$), which is significantly smaller ($p < .01$) than in RECTILINEAR (*mean* $= 21.979$, *sd* $= 28.74$) (Figure 16). We found a significant interaction effect between PLACEMENT and LOCOMOTION ($F_{2,9923.2} = 3.3, p < .05$) where post-hoc pairwise comparison showed that PLACEMENT's effect on CURVILINEAR is significantly different ($p < .05$) from that on RECTILINEAR (Figure 17). We present the signed average in Figure 18.

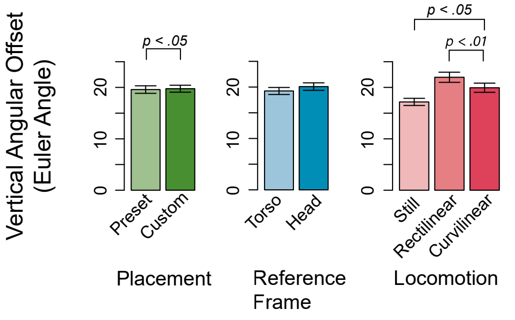


Fig. 16. The absolute vertical angular difference between the target locations and the acquisition locations around participants' bodies in Euler angle degrees.

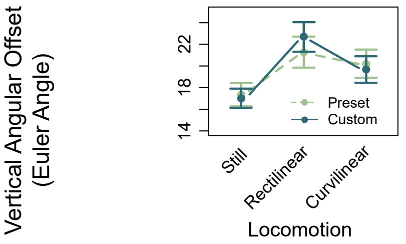


Fig. 17. Two-way interaction effects in VERTICAL ANGULAR OFFSET between PLACEMENT and LOCOMOTION.

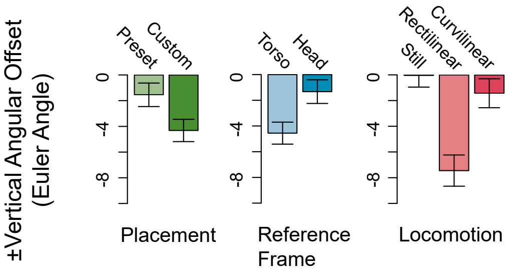


Fig. 18. Signed average of VERTICAL ANGULAR OFFSET: the vertical rotation from the forward vector of the torso to the acquisition location minus the vertical rotation from the forward vector of the torso to the target location.

5 DISCUSSION

In this section, we discuss the statistical results in light of previous work on spatial interaction in immersive MR, on peripersonal space, and on sensory integration of self-motion. We dive into how different reference frames with head or torso, different target placements in relation to the body, and different locomotion patterns collectively affect the spatial perception of invisible targets within the peripersonal space and the acquisition of them in an immersive MR environment.

5.1 Acquisition Error Induced by Walking

From the results, we found that the addition of walking negatively impacted the spatial accuracy of eyes-free target acquisition in immersive MR. This effect is likely due to the body movements generated in the walking dynamics for maintaining gaze- and gait-stability, such as rhythmic oscillations, lateral movements, head turns, and arm swing [10, 29, 33]. Between the two locomotion patterns, RECTILINEAR yielded significantly better performance than CURVILINEAR for the overall EUCLIEAN OFFSET, whereas the opposite was true for VERTICAL ANGULAR OFFSET, while no significant difference was found for HORIZONTAL ANGULAR OFFSET and REACH OFFSET. Walking in a curved path added the turning motion to the complexity of walking in a straight path, and consequently led to more error overall in the acquisition task, as reflected by the difference between the two in EUCLIEAN OFFSET.

For the opposite result found in VERTICAL ANGULAR OFFSET, we believe that it is likely due to the restricted arms swing and associated increase of energy expenditure. During walking and reaching, the reaching motion of the arm is superimposed onto the arm swing naturally generated in the walking motion for compensating the forward stepping motion [7]. Previous works have found that stride length and walking velocity are decreased to maintain gait stability when the arm swing is manually restricted [15]. Additionally, previous works have found that arm swing during bipedal gait serves to reduce the energy expenditure of the whole-body walking motion by facilitating the forward movement of the torso and of the lower limbs [33]. In our study, participants could not slow down their stepping motion because they had to keep up with following the cue, which moved at a fixed velocity. Because participants had to spend more energy to maintain the stability of their gait patterns (i.e., walking velocity and stride length) without the help of arm swing, less energy was preserved for the reaching motion of the arm. As a result, the end points of the acquisition hands were much lower in RECTILINEAR as it induced more forward velocity, which competed for energy with the action of raising the arm.

This is further evidenced by the reduced effect of the restricted arm swing in the CURVILINEAR condition, as the forward velocity was decreased for additional angular velocity towards the walking direction [10]. Our observation of the signed average of VERTICAL ANGULAR OFFSET supports our argument, as the acquisition locations were observably much lower than target locations in RECTILINEAR more than the other locomotion patterns (Figure 18). Further, the significant interaction between PLACEMENT and LOCOMOTION indicated that this difference is more in CUSTOM than in PRESET (Figure 17). We argue that this is due to the benefit from proprioception in PRESET with less bending of the reaching arm. We discuss the effect of PLACEMENT further in the following section.

5.2 Effect of Target Placement in Near vs Far PPS

Previous works have found that spatial perception in the PPS is more accurate for targets closer to the body and that more extreme joint positions get overestimated [42, 53]. However, another study in VR found that target acquisition was more accurate when the targets were placed at arm's length [32]. In this paper, we investigated this question by comparing the performance of eyes-free target acquisition between two conditions where participants were either instructed to learn and grasp invisible targets placed at their arms' lengths (PRESET), or were allowed to place the targets by themselves at more comfortable and easier locations within a small buffer zone during the practice round (CUSTOM) (Figure 5).

We found a statistically significant effect of PLACEMENT in the measures of REACH OFFSET, HORIZONTAL ANGULAR OFFSET, and VERTICAL ANGULAR OFFSET: performance in PRESET was always better than CUSTOM. To interpret these results, we start with REACH OFFSET, the absolute difference between target-torso distance and acquisition-torso distance. We expected participants to place the targets closer to their bodies in CUSTOM when they had the agency to do so, since the targets were forced to be placed at their arms' lengths in PRESET when their arms were stretched out, and they could not have reached further away from their bodies. We can infer this trend from the directional average reach distance in Figure 11, which indicated that the acquisitions were taken at locations closer to the body in CUSTOM much more than in PRESET.

Reflecting upon previous works, we argue that the performance was better in PRESET than in CUSTOM because the stretched-out arms during the reaching motions caused less bending of the joints in the former condition than in the latter. We argue that previous evidence, which are in favour of better performance closer to the body, are less applicable in the case of performing target acquisition during walking in 3D space, because those experiments were conducted on a 2D tabletop while in a seated context. On the other hand, the evidence from the work of Lubos et al. [32], which found more accurate target acquisition at arm's length, was conducted in a similar setting of immersive VR, despite the absence of locomotion. They found that manually reaching towards and acquiring targets at reaching distance with stretched arms turns accurately locating targets from a 3D-complexity task to a 2D-complexity task by minimising the bending of joints. Previous works have also found that active elbow movement generates more disruption of proprioception due to the extra noise in the sensory-motor signal, which leads to an overestimation of arm position [23]. We argue that in the context of walking and reaching in a 3D environment, eyes-free acquisition error in the depth dimension between the body and the target is minimised when targets are located at reaching distance because less disruption of proprioception is induced by elbow movement, which is minimised because of the stretched-out arms. Similarly, we argue that this benefit is also reflected by the accuracy in height. The interaction between PLACEMENT and LOCOMOTION in VERTICAL ANGULAR OFFSET indicated that while the restricted arm swing potentially hindered performance in RECTILINEAR comparing with CURVILINEAR, it is less so in PRESET than in CUSTOM. We argue that this is another result of the less disturbed proprioception overcoming the negative impact on the acquisition performance.

5.3 Effect of Using Head or Torso as Reference Frame

Previous works have found that during walking and turning, the head turns towards the inner part of the walking path more than the torso does for as much as 25° [10, 29]. This discrepancy in the head and torso orientation posits a problem for perceiving the PPS during walking in an MR environment wearing an immersive HMD. In this context, users are only able to perceive visual information rendered through the HMD with a lower temporal-spatial fidelity and within a limited FoV, essentially separating the perception of vision and proprioception into the virtual and the real environments. For instance, popular immersive headsets HTC Vive and Oculus Rift share the same display resolution of $1080\,\text{px} \times 1200\,\text{px}$ per eye, the frame rate of 90 Hz, and the FoV of 110° [6], while human eyes are known to have a horizontal FoV of more than 200° [45]. These limitations in the visual fidelity and in the FoV indicate that users are not able to see their torsos or their hands when they reach into the space out of their FoV, and that self-motion perception becomes less reliable due to limited temporal-spatial resolution for generating optic flow [49].

This discrepancy, as induced by the walking motions, affects the target acquisition actions performed by the arms and hands concurrently. In order to accurately locate targets surrounding our bodies within a reachable distance, we need a reliable mental representation of the peripersonal space, where acquisition and manipulation can be performed [28, 43, 44]. Previous works have found that in different contexts, we adapt the representation of the PPS by using different parts of the bodies as the spatial reference frame depending on the reliability of their associated sensory information [22, 28]. As walking with an immersive HMD induces the separation between visual and proprioceptive sensory information, it was in our interest to investigate how they contribute to the accuracy of the concurrent spatial target acquisition where the brain adaptively integrates the sensory information coming from the separate modalities to form a representation of the out-of-sight targets in the PPS using the head or the torso as the reference frame.

We found that for EUCLIEAN OFFSET and HORIZONTAL ANGULAR OFFSET, using the torso as the reference frame was significantly better than using the head, as the two conditions differ in the horizontal rotation of the targets. We argue that participants were able to locate the targets more accurately by relying on proprioception from their moving hands and arms to determine the orientation of their torsos than relying on head-rotation associated change in vision to determine the orientation of the head during concurrent rotational movements of head and torso. This is in line with ample evidence in previous works that proprioception dominates optic flow when they are in conflict for perception of walking/turning speed, distance travelled, body orientation, and spatial interaction in real and immersive virtual environments [1, 18, 25, 30, 39]. And consequently, this difference in accuracy was reflected in the construction of the representation of the PPS, which led to the different performances of target acquisition. However, we should not overlook the fact that despite initially coming from separate modalities, visual and proprioceptive sensory information end up integrated in the brain for the representation of the PPS. Previous works have found that humans integrate visual and proprioceptive sensory feedback for spatial perception while following a weighted model that takes into account the relative reliability of the different sensory modalities [20, 35, 36]. Using this sensory information, we construct our representation of the PPS to determine out-of-sight arm locations [21]. Whereas proprioception provided more benefit than vision overall, the difference is not as clear-cut when the effect of sensory integration interacts with other factors of eyes-free target acquisition during walking, which we discuss in the following section.

---

[6] https://benchmarks.ul.com/compare/best-vr-headsets

5.4 Interaction between Sensory Integration and Walking Dynamics in the Construction of the PPS

We found a significant main effect of PLACEMENT in all measures but the overall EUCLIEAN OFFSET, where we found a significant two-way interaction effect between PLACEMENT and LOCOMOTION (Figure 8). The significant difference occurred during CURVILINEAR where performance in CUSTOM was better than PRESET more than during the other locomotion patterns. We argue that this is due to the increased radius of the rotational movement of the grasping hand around the body. As participants walked in a curved path, the rotational body movement induced by the turning motion caused more errors in the acquisition task, as the reaching hands rotated around the body. Since we found that the acquisition locations were further away from the body in PRESET due to the stretched-out arms, this increased acquisition-body distance contributed an increased radius of the hand movements following the rotational body movements in curvilinear walking, hence inducing larger error.

We found another two-way interaction in EUCLIEAN OFFSET between REFERENCE FRAME and PLACEMENT where the performance in TORSO was better than in HEAD during PRESET more than during CUSTOM (Figure 8). As discussed previously, better performance in PRESET was due to the undisturbed proprioception for locating out-of-sight targets compared with the noise in the proprioceptive signal generated by the bending joints in CUSTOM. Similarly, performance in TORSO was better because users were relying on proprioception more than in HEAD where the visual perception of the virtual environment and optic flow interferes with the accurate spatial perception provided by proprioception. Consequently, we interpret this two-way interaction as caused by the fact that the undisturbed proprioception was maximally preserved during PRESET + TORSO. In contrast, the combined disturbance from the bending arm joints and the virtual vision peaked during CUSTOM + HEAD.

For HORIZONTAL ANGULAR OFFSET, we found a two-way interaction between REFERENCE FRAME and LOCOMOTION, where performance was better in TORSO than in HEAD during CURVILINEAR more than during STILL (Figure 13). In line with previous discussions, we argue that the difference in performance between HEAD and TORSO was larger during walking in a curved path was because that is when the largest discrepancy between head and torso orientation naturally occurs as supported by previous work [10, 29].

We found two interaction effects between all three factors on EUCLIDEAN and HORIZONTAL ANGULAR OFFSET (Figure 9 and 14). We interpret them together because the significance in both measures was found in the locomotion pattern of RECTILINEAR, where performance was better in TORSO than in HEAD when the targets were placed at maximum reaching distance in PRESET, but not in CUSTOM. In line with previous discussions, we interpret these three-way interactions in light of the literature on sensory integration. Whereas significant difference was mostly found in CURVILINEAR in previously discussed interaction effects, we found significantly different performance during walking in rectilinear paths from the other locomotion patterns in these three-way interactions. Previous results suggested that the discrepancy between the head and torso orientation was induced mostly during curvilinear walking, not during rectilinear walking, which brings us to the question of why performance was better when using the torso as reference frame than when using the head during rectilinear walking.

Without the amplified rotational head movement, we argue that the difference in performance is most likely induced by the integration of visual and proprioceptive sensory information separately perceived from the virtual environment and from the real bodily senses, respectively. Whereas participants walked continuously during CURVILINEAR, their walking path during RECTILINEAR was back-and-forth between two standing points at the boundary of the movement space (Figure 3). The back-and-forth walking induced temporary stopping at the two standing positions where they turned around and walked back. This temporary halting meant that the walking motion was accelerated after each time participants turned around and started walking again. Previous works have found that optic flow is more accurate for the perception of forward motion in accelerated locomotion than in locomotion with slow and constant velocity, during which the forward motion is overestimated for as much as 170% [41].

The accelerated forward motion during RECTILINEAR could have made participants more reliant on vision in the sensory integration process, as more weight would be given to optic flow due to its increased reliability [20, 35, 36]. However, whereas this increased reliability is presumed by the sensory integration process due to the onset of acceleration, the optic flow itself may not be accurate because it was perceived through the HMD which is limited by its resolution and FoV. On the other hand, humans still primarily rely on proprioception, especially in PRESET, to accurately locate targets out-of-sight despite sensory conflict [25]. Consequently, due to the increased disturbance from optic flow when using the head as the reference frame during RECTILINEAR, the performance would enjoy less benefit from proprioception. Thus, we see that RECTILINEAR caused the most sensory conflict in the condition of HEAD + PRESET where optic flow and proprioception had their respective strongest effect, consequently yielding worse performance. Future works can delve deeper and investigate how the sensory integration for forward motion perception is affected by wearing an immersive HMD.

6 DESIGN GUIDELINES

In light of our findings and previous works, we provide the following design guidelines for eyes-free target acquisition in future immersive MR experience featuring spatial mapping and full-body motion tracking. We found that the body movements induced by walking negatively affect the accuracy of eyes-free target acquisition. **GL1: We recommend leaving a buffer zone at the space around the exact locations of the targets to tolerate errors.** For example, consider an imaginary interface for room-scale blueprint editing, as illustrated in Figure 1. It enables its users to walk within the virtual environment as a 3D canvas with spatial mapping to a physical room, to perform fine-grained selection and manipulation of elements of the blueprint. Users acquire tools such as a magnifier by grabbing within the virtual sphere following their bodies. Following **GL1**, the spheres need to become larger when walking motion of the user is detected to tolerate potentially larger acquisition errors.

We also found that, as a consequence of restricting arm swing, the acquisition locations tend to be lower than the targets actual location during forward locomotion. While the error in the height of the target acquisitions during walking is larger than when standing still, this error increases as the curvature of the walking path decreases. **GL2: Future designs should leave more error tolerance along the vertical axis of the targets for acquisition tasks during forward locomotion.** In the case of the blueprint-editing interface, more error tolerance (enlargement) should be added to the area below the actual spheres when users perform acquisition actions while walking forward.

As the curvature of the walking path increases, the concurrent walking and turning induce a discrepancy between head and torso orientation. We found that the eyes-free target acquisition performance is better when using the torso as the reference frame to perceive the target locations in the space around our bodies. **GL3: We advise future systems to provide torso position tracking rather than the traditional head tracking function whenever possible.** As an example, the tools in the blueprint-editing interface should follow the orientation of trackers attached on the torso instead of the HMDs of the users. This will provide a more reliable spatial reference frame for eyes-free target acquisition and for users to perceive the locations of virtual content around their bodies in general.

Finally, we found that placing targets at users' reaching distances helps improve eyes-free target acquisition performance due to less bending of the reaching arm and the consequent greater benefit from accurate proprioception. **GL4: Future designs should arrange the targets at different angular locations around the body at the fixed distance of the approximate arm's length to minimise target acquisition error.** In the case of the blueprint-editing interface, the tool spheres should be anchored around the user's body at the approximate maximum reaching distance.

7 LIMITATIONS AND FUTURE WORK

To investigate users' undisturbed performance eyes-free target acquisition in immersive MR, we did not provide feedback on their accuracy during the study. Additionally, we designed the study to be free of context to minimise extra cognitive load induced by the tasks. The use of intuitive hand gestures without arbitrary feedback makes our work instructive for a wide range of future works that evaluate similar interaction with variations in software and hardware settings. Future works can investigate the effect of different types of feedback in different contexts, such as games and utilitarian applications, and collect subjective feedback to aid the understanding of the performance results. We used the grasping gesture to trigger the acquisition instead of using pointing with fingers or with controllers to simulate the natural interaction that we perform with real-world objects [5]. Future studies need to take the potential effect of gesture into account if they are limited by using controllers or certain gesture presets.

While we designed the study to investigate the potential factors affecting target acquisition in the PPS out of the FoV, we did not compare the performance of acquiring targets placed at different locations around the body. Previous works have investigated the performance of eyes-free target acquisition under the effect of target placement at many different angles in VR while sitting or standing still [55]. Future works can investigate the effect of different layouts of targets inside and outside of the FoV, with participants wearing immersive and see-through headsets. We did not expect any effect from hand dominance because we designed the study to make participants perform the task using their left and right hands for the equal number of times due to the symmetrical target allocation. However, future works can investigate this effect to potentially help improve user experience. While we designed the target distribution out of the FoV of most common HMDs, variations in resolution and brightness of different HMDs may still affect the availability and intensity of optic flow. Future works should take this into consideration.

In this study, we investigated the different effects of rectilinear and curvilinear walking by comparing straight walking motion against circular motion, with a constant movement velocity. It would be interesting to explore how different path curvatures with smaller intervals (such as 10°) affect the performance differently under different walking speeds in future work. In addition, future works can investigate the potential effect of cognitive load induced by walking in more complex patterns while acquiring targets in specific application contexts. Similarly, we investigated the performance under two conditions with PLACEMENT: when targets placed at reaching distance and when a small buffer was allowed for participants to place the targets closer to their bodies. Future works can investigate how different target-body distances with smaller intervals (such as 0.1m) affect the performance. Additionally, we applied the same walking speed, path curvatures, and the buffer size of target placement in CUSTOM across participants. Future works can expand the investigation to cover the effect of individual differences such as height, arm length, and gait length.

Finally, we proposed the design guidelines based on our interpretation of the statistical findings obtained in a controlled lab setting with close reference to previous works, However, our arguments and speculations for the causes of the significant results are likely not exhaustive, as our findings may have been partly induced by hidden factors other than those that we investigated and discussed. Our design guidelines apply primarily to future works in which the factors we investigated are important to the user experience.

8 CONCLUSION

In this paper, we investigated user performance of eyes-free target acquisition during walking in immersive Mixed Reality with spatial mapping to the real environment. In a study with 24 participants, we measured the performance of grabbing targets located outside of the FoV of an immersive HMD. We measured the performance in the form of the Euclidean and angular offsets and the errors in reaching distance between the actual target locations and the locations of users' grasping hands. We compared performance with different target placement mechanisms, with head or torso as reference frames, and in different locomotion conditions where participants stood still, walked in rectilinear paths, or walked in curvilinear paths.

The results showed that the performance was significantly better when targets followed users' torso orientations than their head orientations. We also found that the errors of the angular rotations and the reaching distances of the reaching arms were smaller when the targets were placed at users' approximate arm lengths while benefiting from less bending of the arms. The overall performance also demonstrated that target acquisition was more challenging for users when they walked in the immersive MR environment than when standing still, especially during curvilinear locomotion. We discussed other findings in light of related works on sensory integration and self-motion perception. With this paper, we contribute the evaluation of an important daily task, eye-free target acquisition during locomotion, in the context of immersive MR. We contribute understanding of the sensory integration mechanism involved in the representation and the update of the representation of peripersonal space during walking in immersive MR. Finally, with our findings, we provide design guidelines for future immersive MR experience featuring spatial mapping and full-body motion tracking.

REFERENCES

[1] N. H. Bakker, P. J. Werkhoven, and P. O. Passenier. The effects of proprioceptive and visual feedback on geographical orientation in virtual environments. *Presence: Teleoperators and Virtual Environments*, 8(1):36–53, 1999. doi: 10.1162/105474699566035

[2] M. Baloup, T. Pietrzak, and G. Casiez. Raycursor: A 3d pointing facilitation technique based on raycasting. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19, pp. 101:1–101:12. ACM, New York, NY, USA, 2019. doi: 10.1145/3290605.3300331

[3] W. Becker, G. Nasios, S. Raab, and R. Jürgens. Fusion of vestibular and podokinesthetic information during self-turning towards instructed targets. *Experimental Brain Research*, 144(4):458–474, June 2002. doi: 10.1007/s00221-002-1053-5

[4] M. Billinghurst, J. Bowskill, N. Dyer, and J. Morphett. An evaluation of wearable information spaces. In *Proceedings. IEEE 1998 Virtual Reality Annual International Symposium (Cat. No.98CB36180)*, pp. 20–27, March 1998. doi: 10.1109/VRAIS.1998.658418

[5] C. Brozzoli, L. Cardinali, F. Pavani, and A. Farnè. Action-specific remapping of peripersonal space. *Neuropsychologia*, 48(3):796 – 802, 2010. The Sense of Body. doi: 10.1016/j.neuropsychologia.2009.10.009

[6] L. Cardinali, C. Brozzoli, and A. Farnè. Peripersonal Space and Body Schema: Two Labels for the Same Concept? *Brain Topography*, 21(3):252–260, May 2009. doi: 10.1007/s10548-009-0092-7

[7] H. Carnahan, B. J. Mcfadyen, D. L. Cockell, and A. H. Halverson. The combined control of locomotion and prehension. *Neuroscience Research Communications*, 19(2):91–100, 1996. doi: 10.1002/(SICI)1520-6769(199609)19:2<91::AID-NRC168>3.0.CO;2-X

[8] E. Chiovetto and M. A. Giese. Kinematics of the coordination of pointing during locomotion. *PloS one*, 8(11):e79555–e79555, Nov. 2013. Publisher: Public Library of Science. doi: 10.1371/journal.pone.0079555

[9] A. Cockburn, P. Quinn, C. Gutwin, G. Ramos, and J. Looser. Air pointing: Design and evaluation of spatial target acquisition with and without visual feedback. *International Journal of Human-Computer Studies*, 69(6):401 – 414, 2011. doi: 10.1016/j.ijhcs.2011.02.005

[10] G. Courtine and M. Schieppati. Human walking along a curved path. i. body trajectory, segment orientation and the effect of vision. *European Journal of Neuroscience*, 18(1):177–190, 2003. doi: 10.1046/j.1460-9568.2003.02736.x

[11] C. Cruz-Neira, D. Reiners, and J. P. Springer. An affordable surround-screen virtual reality display. *Journal of the Society for Information Display*, 18(10):836–843, 2010. doi: 10.1889/JSID18.10.836

[12] R. P. Darken, W. R. Cockayne, and D. Carmein. The omni-directional treadmill: A locomotion device for virtual worlds. In *Proceedings of the 10th Annual ACM Symposium on User Interface Software and Technology*, UIST '97, pp. 213–221. ACM, New York, NY, USA, 1997. doi: 10.1145/263407.263550

[13] G. di Pellegrino and E. Làdavas. Peripersonal space in the brain. *Neuropsychologia*, 66:126 – 133, 2015. doi: 10.1016/j.neuropsychologia.2014.11.011

[14] F. H. Durgin, A. Pelah, L. F. Fox, J. Lewis, R. Kane, and K. A. Walley. Self-motion perception during locomotor recalibration: more than meets the eye. *Journal of Experimental Psychology: Human Perception and Performance*, 31(3):398, 2005.

[15] S. Eke-Okoro, M. Gregoric, and L. Larsson. Alterations in gait resulting from deliberate changes of arm-swing amplitude and phase. *Clinical Biomechanics*, 12(7):516 – 521, 1997. doi: 10.1016/S0268-0033(97)00050-8

[16] B. M. Ens, R. Finnegan, and P. P. Irani. The personal cockpit: A spatial interface for effective task switching on head-worn displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '14, pp. 3171–3180. ACM, New York, NY, USA, 2014. doi: 10.1145/2556288.2557058

[17] M. O. Ernst and M. S. Banks. Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, 415(6870):429–433, Jan. 2002. doi: 10.1038/415429a

[18] H. Frenz and M. Lappe. Absolute travel distance from optic flow. *Vision Research*, 45(13):1679 – 1692, 2005. doi: 10.1016/j.visres.2004.12.019

[19] I. Frissen, J. L. Campos, J. L. Souman, and M. O. Ernst. Integration of vestibular and proprioceptive signals for spatial updating. *Experimental Brain Research*, 212(2):163, May 2011. doi: 10.1007/s00221-011-2717-9

[20] J. J. Gibson. *The perception of the visual world.* The perception of the visual world. Houghton Mifflin, Oxford, England, 1950. Pages: xii, 242.

[21] M. S. A. Graziano. Where is my arm? The relative role of vision and proprioception in the neuronal representation of limb position. *Proceedings of the National Academy of Sciences*, 96(18):10418–10421, 1999. doi: 10.1073/pnas.96.18.10418

[22] M. S. A. Graziano, X. T. Hu, and C. G. Gross. Visuospatial properties of ventral premotor cortex. *Journal of Neurophysiology*, 77(5):2268–2292, 1997. PMID: 9163357. doi: 10.1152/jn.1997.77.5.2268

[23] V. Gritsenko, N. I. Krouchev, and J. F. Kalaska. Afferent input, efference copy, signal noise, and biases in perception of joint angle during active versus passive elbow movements. *Journal of Neurophysiology*, 98(3):1140–1154, 2007. PMID: 17615137. doi: 10.1152/jn.00162.2007

[24] P. Haggard and D. M. Wolpert. Disorders of body scheme. In *In Freund, H.J., Jeannerod, M., Hallett, M., Leiguarda R., (Eds.), Higher-Order Motor Disorders*. University Press, 2005.

[25] L. R. Harris, M. Jenkin, and D. C. Zikovitz. Visual and non-visual cues in the perception of linear self motion. *Experimental Brain Research*, 135(1):12–21, Nov. 2000. doi: 10.1007/s002210000504

[26] L. J. Hettinger and G. E. Riccio. Visually induced motion sickness in virtual environments. *Presence: Teleoperators and Virtual Environments*, 1(3):306–310, 1992. doi: 10.1162/pres.1992.1.3.306

[27] A. Hilsendeger, S. Brandauer, J. Tolksdorf, and C. Fröhlich. Navigation in virtual reality with the wii balance board. In *6th Workshop on Virtual and Augmented Reality*, 2009.

[28] S. B. Hunley and S. F. Lourenco. What is peripersonal space? an examination of unresolved empirical issues and emerging findings. *WIREs Cognitive Science*, 9(6):e1472, 2018. doi: 10.1002/wcs.1472

[29] T. Imai, S. T. Moore, T. Raphan, and B. Cohen. Interaction of the body, head, and eyes during walking and turning. *Experimental Brain Research*, 136(1):1–18, Jan. 2001. doi: 10.1007/s002210000533

[30] M. J. Kearns, W. H. Warren, A. P. Duchon, and M. J. Tarr. Path integration from optic flow and body senses in a homing task. *Perception*, 31(3):349–374, 2002. PMID: 11954696. doi: 10.1068/p3311

[31] F. C. Y. Li, D. Dearman, and K. N. Truong. Virtual shelves: Interactions with orientation aware devices. In *Proceedings of the 22Nd Annual ACM Symposium on User Interface Software and Technology*, UIST '09, pp. 125–128. ACM, New York, NY, USA, 2009. doi: 10.1145/1622176.1622200

[32] P. Lubos, G. Bruder, O. Ariza, and F. Steinicke. Touching the sphere: Leveraging joint-centered kinespheres for spatial user interaction. In *Proceedings of the 2016 Symposium on Spatial User Interaction*, SUI '16, pp. 13–22. ACM, New York, NY, USA, 2016. doi: 10.1145/2983310.2985753

[33] P. Meyns, S. M. Bruijn, and J. Duysens. The how and why of arm swing during human walking. *Gait & Posture*, 38(4):555 – 562, 2013. doi: 10.1016/j.gaitpost.2013.02.006

[34] M. R. Mine, F. P. Brooks, Jr., and C. H. Sequin. Moving objects in space: Exploiting proprioception in virtual-environment interaction. In *Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '97, pp. 19–26. ACM Press/Addison-Wesley Publishing Co., New York, NY, USA, 1997. doi: 10.1145/258734.258747

[35] H. Mittelstaedt. Interaction of eye-, head-, and trunk-bound information in spatial perception and control. *Journal of vestibular research : equilibrium & orientation*, 7(4):283–302, Aug. 1997. Place: Netherlands.

[36] M.-L. Mittelstaedt and S. Glasauer. Idiothetic navigation in gerbils and humans. *Zool. Jb. Physiol*, 95(427-435):212, 1991.

[37] J.-P. Noel, P. Grivaz, P. Marmaroli, H. Lissek, O. Blanke, and A. Serino. Full body action remapping of peripersonal space: The case of walking. *Neuropsychologia*, 70:375 – 384, 2015. doi: 10.1016/j.neuropsychologia.2014.08.030

[38] J.-P. Noel, A. Serino, and M. T. Wallace. Increased neural strength and reliability to audiovisual stimuli at the boundary of peripersonal space. *Journal of Cognitive Neuroscience*, 31(8):1155–1172, 2019. PMID: 30188779. doi: 10.1162/jocn_a_01334

[39] P. Pretto, M. Ogier, H. Bülthoff, and J.-P. Bresciani. Influence of the size of the field of view on motion perception. *Computers & Graphics*, 33(2):139–146, Apr. 2009. doi: 10.1016/j.cag.2009.01.003

[40] S. Razzaque, D. Swapp, M. Slater, M. C. Whitton, and A. Steed. Redirected walking in place. In *EGVE*, vol. 2, pp. 123–130, 2002.

[41] F. P. Redlick, M. Jenkin, and L. R. Harris. Humans can use optic flow to estimate distance of travel. *Vision Research*, 41(2):213–219, Jan. 2001. doi: 10.1016/S0042-6989(00)00243-1

[42] L. Rincon-Gonzalez, C. A. Buneo, and S. I. Helms Tillery. The proprioceptive map of the arm is systematic and stable, but idiosyncratic. *PloS one*, 6(11):e25214–e25214, 2011. Edition: 2011/11/16 Publisher: Public Library of Science. doi: 10.1371/journal.pone.0025214

[43] G. Rizzolatti, L. Fadiga, L. Fogassi, and V. Gallese. The Space Around Us. *Science*, 277(5323):190–191, 1997. doi: 10.1126/science.277.5323.190

[44] G. Rizzolatti, C. Scandolara, M. Matelli, and M. Gentilucci. Afferent properties of periarcuate neurons in macaque monkeys. I. Somatosensory responses. *Behavioural Brain Research*, 2(2):125 – 146, 1981. doi: 10.1016/0166-4328(81)90052-8

[45] T. C. Ruch and J. F. Fulton. Medical Physiology and Biophysics. *Academic Medicine*, 35(11), 1960.

[46] A. Serino. Peripersonal space (PPS) as a multisensory interface between the individual and the environment, defining the space of the self. *Neuroscience & Biobehavioral Reviews*, 99:138 – 159, 2019. doi: 10.1016/j.neubiorev.2019.01.016

[47] A. Serino, J.-P. Noel, G. Galli, E. Canzoneri, P. Marmaroli, H. Lissek, and O. Blanke. Body part-centered and full body-centered peripersonal space representations. *Scientific Reports*, 5(1):18603, Dec. 2015. doi: 10.1038/srep18603

[48] J. L. Souman, P. R. Giordano, M. Schwaiger, I. Frissen, T. Thümmel, H. Ulbrich, A. D. Luca, H. H. Bülthoff, and M. O. Ernst. Cyberwalk: Enabling unconstrained omnidirectional walking through virtual environments. *ACM Trans. Appl. Percept.*, 8(4):25:1–25:22, Dec. 2008. doi: 10.1145/2043603.2043607

[49] F. Steinicke, Y. Visell, J. Campos, and A. Lcuyer. *Human Walking in Virtual Environments: Perception, Technology, and Applications*. Springer Publishing Company, Incorporated, 2013.

[50] J. L. Taylor. Proprioception. In L. R. Squire, ed., *Encyclopedia of Neuroscience*, pp. 1143 – 1149. Academic Press, Oxford, 2009. doi: 10.1016/B978-008045046-9.01907-0

[51] B. R. Umberger. Effects of suppressing arm swing on kinematics, kinetics, and energetics of human walking. *Journal of Biomechanics*, 41(11):2575 – 2580, 2008. doi: 10.1016/j.jbiomech.2008.05.024

[52] B. A. Whitehead. James j. gibson: The ecological approach to visual perception. boston: Houghton mifflin, 1979, 332 pp. *Behavioral Science*, 26(3):308–309, 1981. doi: 10.1002/bs.3830260313

[53] E. T. Wilson, J. Wong, and P. L. Gribble. Mapping proprioception across a 2D horizontal workspace. *PloS one*, 5(7):e11851–e11851, July 2010. Publisher: Public Library of Science. doi: 10.1371/journal.pone.0011851

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

[55] Y. Yan, C. Yu, X. Ma, S. Huang, H. Iqbal, and Y. Shi. Eyes-free target acquisition in interaction space around the body for virtual reality. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, pp. 42:1–42:13. ACM, New York, NY, USA, 2018. doi: 10.1145/3173574.3173616