# Walking on Powered VR Shoes to Virtual Reality Motion: a User Experience Evaluation

## Master thesis by Amber Elferink

Utrecht University student number: 5491525
TU Delft student number: 5418526

## Supervision by

Dr. ing. L. Marchal-Crespo
Dr. W.O. Hürst
Dr. A.F.F. Derumigny
A. Foxcroft

TU Delft
Mechanical Engineering (3ME)

Utrecht University
Faculty of Science

Assesment Committee
Dr. ing. L. Marchal-Crespo
Dr. A.F.F. Derumigny
External: G. Papaioannou
External: M.L. van de Ruit

Assesment Committee
Dr. W.O. Hürst
External: Prof. dr. R.C. Veltkamp

# CONTENTS

# Walking on Powered VR Shoes to Virtual Reality Motion: a User Experience Evaluation

Amber Elferink

***Abstract*—Moving through immersive virtual reality (VR) is commonly achieved by physically walking in the real room or using other techniques like an omnidirectional treadmill or walk-in-place. Roomscale walking is most similar to normal walking but is limited by physical space. However, other techniques can cause user experience issues such as VR sickness, balance problems, and feeling unnatural. Newer locomotion techniques are available such as powered VR shoes, which are shoes with motorized treadmills underneath. While walking, the shoes drive the user backward and actively negate the forward velocity, reducing the needed physical space. Yet, there is little evidence of the effect of powered VR shoes on user experience, which part of this work addresses. Additionally, previous research shows mismatched VR motion (optical flow) can increase VR sickness, cognitive load, and break presence. However, full-gait locomotion studies often focus on the device, neglecting optical flow, and what is the best body part to control optical flow direction is still an open question. Therefore, we first developed a novel algorithm to convert leg-based walking to optical flow while walking on VR shoes, which may also be used for other full-gait locomotion techniques. We conducted a study with 20 participants to find which of four optical flow implementations, differing in VR motion direction, resulted in the best user experience. These direction conditions were based on body-mounted trackers: i) head orientation, ii) hip orientation, iii) standing foot velocity direction, and iv) average orientation of both feet. Head-oriented walking resulted in a significantly worse user experience compared to other conditions, with no significant differences among any other conditions. Additionally, we found no effect of optical flow on VR sickness, Mental Effort, and Presence, contrary to previous studies, but instead significant differences in Ease of Use, Input responsiveness, and Appropriateness, and indication that other user experience factors might be impacted more. Finally, we discovered that walking on VR shoes, although not completely comfortable and natural, was learnable within 10 minutes for all participants under 60 years old.**

***Index Terms*—virtual reality, locomotion, walking, user experience, tracking, direction, steering, shoes, feet, hip, head, powered, active, motorized, gait, transfer function, optical flow, redirected walking, omnidirectional treadmill**

## I. INTRODUCTION

Walking in virtual worlds is widely used. Some examples of applications are training in dangerous situations [1], rehabilitation research [2], research to remotely control robots [3], and games. Immersive Virtual Reality (VR) allows users to experience virtual worlds through head-mounted displays. Moving through these virtual worlds (locomotion) in VR can be done in multiple ways [4]. The most similar VR locomotion technique to normal walking is walking at room-scale, with the VR headset showing the virtual world moving 1:1 relative to the user, as would be expected while walking in the physical world. However, with large to infinite VR worlds, the limited size of the physical room becomes a problem. Multiple locomotion solutions have tried to mitigate this. While the real world has physical constraints, VR does not. This flexibility enables any conceivable mapping from physical movements to virtual ones, such as joystick motion, walking-in-place, and many more, see subsection I-A for a summary.

However, these techniques introduce new user experience (UX) issues like VR sickness [5]–[8], the devices feeling unnatural [7]–[12] or causing balance problems [13], being tiring [7], [14], being more difficult, less enjoyable, than walking on a regular treadmill [7], lack of control [15] or reduction of presence in the virtual world [16]. Finding a locomotion method that mitigates user experience problems like these, while only requiring a small physical space, is one of VR's primary challenges [4], [17].

VR sickness is especially uncomfortable since it can result in symptoms like nausea, similar to motion sickness symptoms experienced in a car or boat. According to the sensory conflict theory, this is caused by a mismatch between optical flow (apparent visual motion) and vestibular queues [18]. For example, using a joystick to smoothly move in VR can lead to more VR sickness compared to physically walking [19]. Thereby, this optical flow mismatch not only can introduce VR sickness [20], [21], but also can also increase cognitive load [21], and break presence, which means the user no longer feels like they are in the virtual environment [16], even when the optical flow mismatch is introduced while walking in room-scale.

We focused on evaluating a new promising locomotion technique: powered VR shoes that drive the user back while walking. This is a novel technique, which to our knowledge, has not been evaluated in a user study, and therefore its user experience limitations are currently unknown.

As we discussed, optical flow plays a major role in VR sickness and other user experience problems. Previous full-gait techniques often focus on the device, neglecting the optical flow, which causes little to be known about how this optical flow should be matched. It remains an open question which body part to rely on for optical flow direction [22]. Therefore, we developed several options for calculating virtual optical flow based on tracker motion based on head, hip, and feet and compared those in user experience. We also expected, and found, that optical flow influences more user experience factors than the few measured in previous works. In summary, we investigate:

- Q1) Which of our optical flow mappings, specifically evaluating direction, gives the best user experience while walking on VR shoes?
- Q2) What user experience factors are most influenced by differences in optical flow?

- Q3) What are the current user experience limitations of VR shoes?

We discuss our contributions in more detail in subsection I-C. First, we will discuss different leg-based locomotion techniques and their user experience limitations subsection I-A, then previous works on the transfer function/mapping (the algorithm converting sensor data to optical flow) in subsection I-B. In our methods, we discuss the VR shoe device in subsection II-A, followed by our transfer function algorithm, subsection II-B. Then we describe our user test in subsection II-C, followed by the statistical analysis subsection II-D, which gave the results in section III. We discuss these results in section IV. Appendix A discusses how we came to our speed algorithm, including some analyses and suggestions for improvement. Appendix B to Appendix G, is organized based on the measures, with for each of these the hypotheses, statistics methods, results, and discussion briefly summarized. In Appendix H the full questionnaire is shown.

### A. Locomotion techniques

Locomotion techniques that mimic walking can be divided between full gait techniques where the user goes through the full gait cycle, and a partial gait technique: Walk-in-place (WIP) where the user makes a marching or heel-tapping motion on the same spot. With data from trackers mounted to their feet, this motion can be converted to a virtual motion velocity. Full gait techniques can be further divided into redirection, passive, and active locomotion techniques. Redirection techniques involve room-scale walking, but with an additional virtual illusion that rotates or scales movements in the virtual world. This virtual illusion can cause the user to walk in a curve while giving the feeling of walking straight. Applying this strategically, the physical space required can be much smaller while walking a large area virtually. However, the minimum circle size required for such a technique must be between 6.5 to 22 m radius, to cause only 50% of people not to realize they are walking a circle [23]. Additionally, VR sickness is still often a problem [20], [24].

We can also find multiple passive motion negation devices: slidemills, passive VR shoes, and a human-sized hamster ball. Slidemills are large devices that hold the user around their waist while providing a slippery surface beneath the feet. This way the user can make a walking or running motion without moving in the physical world since they are restrained by the hip band, which also registers their direction of walking. Although this movement looks similar to walking, users do not feel it is similar to walking. Walking on a slidemill is often perceived as more fatiguing and builds more strain compared to walking and other locomotion techniques [7]–[11]. Additionally, slidemills do not reduce VR sickness, and come out of user testing as the worst option on most criteria compared to WIP in most studies found [6]–[8]. The human-sized hamster ball [25] has similar problems to the slidemills, with the additional issue of starting and stopping the spheres momentum [26], [27].

Active devices can be subdivided into omnidirectional treadmills, active VR shoes, and some other devices as exoskeletons [28], [29] and moving robotic tiles[30], [31]. Omnidirectional treadmill [32]–[36] is a name often used for slidemills in literature. However, we distinguish them here since true omnidirectional treadmills, similar to a conventional treadmill, have a surface that moves powered by motors. However, omnidirectional treadmills can be actuated in any horizontal direction. This allows users to achieve a more similar biomechanical symmetry, which means that movements and forces are similar to normal walking [37]. To control the treadmill surface, a tracker is attached to the user's hip height at the back, and the treadmill is actuated to keep it in the center of the treadmill via more or less complex control algorithms [13], [32]. However, especially when walking corners and during stopping, users currently have difficulties maintaining balance [13], [38], even with state-of-the-art devices [32], [39]. This is mostly due to the responsiveness of the device due to inertia and the control algorithm. Recent developments, for example, Pyo, Lee, and Yoon [32], are capable of accelerating fast enough to counter the former. Therefore, the difficulty lies mostly in developing control algorithms that do make the surface respond as we would experience the real floor moving with respect to the user.

VR shoes are a relatively new technique, and thus, user tests are not found in research so far. There are passive VR shoes, meaning they contain no motors and have free rolling rollers similar as on roller skates [40], [41], and active VR shoes, meaning motor driven without free rolling [42]–[44]. Cybershoes [41] are passive commercial shoe-like devices bound under the user's feet with a single roller per foot that detects rolling over the ground. These are not intended to stand on while walking, but instead, the user sits on a rotating chair, moving their feet along the ground as if walking. Finally Functional [40] shows DIY prototypes for passive VR shoes where the user is standing and constrained by the hips similar to a slidemill.

For active VR shoes, Iwata, Yano, and Tomioka [44] describes a VR shoes prototype with a backpack, and Google describes active VR shoes in a patent [45]. Finally Functional [46] also created a prototype for large active VR shoes, with a base that stays on the ground while the foot can be lifted from it. EktoVR [42] is a company that developed VR boots on wheels powered by motors. Their control algorithm gradually activates and deactivates the shoe's movement when starting or stopping to walk, which still requires a large playspace.

Despite their potential impact in improving VR locomotion, currently, it is not known how users experience VR shoes. None of the VR shoes found in the literature have undergone user testing, aside from a journalist who described his own experience with the Ekto-VR shoes [47]. Therefore, in this paper, we evaluated the user experience with the early prototype of active VR shoes from Freeaim [43]. These VR shoes are wireless, do not need a backpack and the total device size is smaller than the other active VR shoes found in literature. The prototype used here can only move the users' feet from front to back (no side-stepping). The shoes try to quickly respond to starting and stopping, aiming for a playspace of 2x2 m.

Little research has been done on the user experience of VR shoes. Aside from the locomotion techniques themselves, the transfer function (how to map physical motion to virtual motion) is also an important part of a locomotion experience, and will be compared in this paper. Therefore, in the next section, we first discuss previous works on transfer functions for locomotion techniques mimicking real walking.

### B. Previous works on the transfer function in walking-related locomotion

The transfer function is the mapping from physical motion to virtual motion and consists of two main parts: translation and direction [48]. These two are then combined to obtain a virtual velocity.

*1) Speed/Translation:* Works using older VR headsets and researching the match of the physical treadmill and VR speed have found that the virtual speed is perceived as too slow compared to the real walking speed [49]–[51]. However, this seems to be due to the limited field-of-view (FOV) of the VR headsets used and does not seem to be present anymore with modern VR headsets with FOV above 110 degrees diagonal [52]. Steinicke, Bruder, Jerald, *et al.* [53] found that the detection threshold (50% of users notice it) for up- or downscaling virtual speed is a factor of 0.86 to 1.26 from users' physical walking speed. However, it is unclear if this is still the case for modern VR headsets with larger FOV.

Since we focus on analyzing direction in this work, previous works on the optical flow algorithms of speed are discussed in subsection A-A in Appendix A.

*2) Directon/Steering:* In real walking without a headset, the head orientation anticipates the future direction of walking by 200 to 600 ms before the body turns. Even when constraining the head movement concerning the trunk, the trunk aligns faster so the head can align with the future direction [54]. The gaze looks ahead further than the head and seems to initiate the in-sequence turning order of the head, trunk, hip, and feet. If the gaze is fixed, all body parts are turned in unison, discarding the sequential order in a test with WIP [54].

Nilsson, Serafin, and Nordahl [22] suggests that what body part to rely on for the user's virtual walking direction is still an open question, and encourages future work to look into this. This would probably also depend on the locomotion technique used. This the user's virtual walking direction is important to consider, since even in walking on normal ground, a mismatch in this direction optical flow can result in VR sickness [20], [21], and increase cognitive load [21]. Additionally, these differences in optical flow have to stay under the detection thresholds, or more than 50% of people will notice, which may break presence [16]. These detection thresholds have been researched extensively in redirected walking [23], which in essence is normal walking with a mismatching optical flow. However, most works researching this still use outdated VR headsets with small FOVs, so it is unclear if this holds up with modern headsets. Even under the detection thresholds, redirected walking can cause VR sickness [20]. Therefore, it is important to present a matching optical flow to the user. In this work, we investigate the open question of which body part to rely on for the user's virtual walking direction while walking on active VR shoes.

Virtual walking direction mappings have often been researched for the WIP technique. Al Zayer, MacNeilage, and Folmer [55] gives an overview of WIP works and the different methods they used to control speed and direction. Direction was controlled using head, knees, torso, hip, and hands with yawing, tilting, or leaning. Additionally Nilsson, Serafin, Laursen, *et al.* [56] used the average orientation of the feet as direction.

An advantage of head-directed steering is that no other trackers are needed. However, it limits the user's ability to look around the environment without unwillingly changing direction [57]. When using hands, it deprives the user of proprioceptive and kinesthetic feedback from whole body turns [57], which in turn reduces spatial awareness [38], [58]. The head direction was found better compared to the torso in spatial orientation and user preference during WIP. [59]. Tan, Foo, Yeo, *et al.* [15] also found their head-oriented technique better in perceived naturalness compared to their WIP technique using torso orientation and average hand orientation, although the gestures themselves also differed between conditions and therefore could also have impacted this finding. Williams, McCaleb, Strachan, *et al.* [59] theorize that the reduced torso preference and spatial orientation compared to the head direction could have been caused by it being disorienting for users to look around while walking. However, in their experiment, users first walked and then turned, so they were never turning while walking in place. When comparing torso steering to lower body steering, torso steering has been suggested by authors to be less natural and cause more undesirable rotations than hip or foot-directed steering [22], [60], but this was not evaluated in a user test.

Rebenitsch and Engle [61] tested preference between semi-decoupled head, and regular head oriented and hand oriented. In the semi-decoupled method, the joystick was used for body direction, and the VR headset was used for the head. When the head turned the body would gradually realign over a few seconds. The semi-decoupled method was never preferred over the other conditions; 68% of participants chose hand-oriented steering while the rest preferred the head condition.

### C. Contributions of this work

As we discussed earlier, matching optical flow is important to reduce VR sickness [20], [21], reduce cognitive load [21], and maintain presence [16]. While important, so far this topic has been neglected in current full-gait techniques, and an optical flow algorithm from full-gait physical walking motion on a user-centering locomotion device has as far as we know not been described.

Therefore, we present an algorithm to convert this physical movement to virtual movement from foot trackers (and optionally hip and head trackers), which we evaluate for walking on VR shoes, and may also be suitable for other locomotion techniques with full gait. There are two components to this virtual movement: speed and direction. Since the speed factor seems researched by some previous works and can be done with a normal treadmill, our work focuses on the direction component. Our research question is: *Which orientation mapping, choosing from headset, standing foot, average shoes,* *or hip orientation to VR optical flow gives the best user experience while walking on VR shoes?*

Thereby, we expect optical flow user experience problems are not limited to the VR sickness, cognitive load, and presence tested in previous works. Therefore we investigate a wide range of user experience factors often researched in locomotion to see which are most impacted by differences in optical flow. Many works researching VR locomotion user experience only measure a few attributes of user experience, such as presence, VR sickness, cognitive load, or usability. Cannavò, Calandra, Pratticò, *et al.* [14] created a standardized testbed with multiple virtual scenarios and collected 20 user experience attributes, combining them into a single questionnaire. This work will be used as the basis for the user experience testing. This questionnaire and our adjustments are further discussed in subsubsection II-C5. We compare the four different virtual direction options from our virtual movement algorithm, while the user walks on powered VR shoes. The raw data, including individual tracker data that can be replayed in Unity, is freely published, see Github [62].

Thereby, to our knowledge, VR shoes were not publicly evaluated on user experience, which means we do not know if they are a promising step toward solving the locomotion problem. We assess the VR shoes current limitations in this user test.

### D. Choice of conditions and hypotheses

The head direction condition was chosen because it is the default option in most applications, it is the most often tested option in literature, and it requires no extra trackers. Controller/hand orientation is not tested, since they occupy the hands preventing handling items during walking, and are less related to our automatic natural body orientation. The hip was chosen over the torso position due to the suggestions of reducing undesirable rotations and increasing naturalness [22], [60]. The standing foot velocity direction was implemented because in normal walking the standing foot pushes the body in the opposite direction of the standing foot velocity. Similarly, the standing foot orientation is used for Cybershoes [41]. The average feet orientation was proposed since the lifted foot might slightly anticipate the direction of the curve. Therefore, combining this with the standing foot could give a more responsive virtual direction. Additionally, this method was the only foot-based direction used for WIP found in literature [56].

Corresponding to the three main topics investigated, the hypotheses are:

- Hypotheses Q1: the standing foot velocity direction gives the best user experience since, in theory, it is most similar to normal walking. The head gives the worst user experience since the head direction orients 200 to 600 ms further ahead of the curve, and is not decoupled.
- Hypotheses Q2: many user experience factors are influenced by optical flow differences, and VR sickness is the most influenced.
- Hypotheses Q3: users have to learn to walk on the VR shoes and may have problems with balance at the start since the standing foot velocity with respect to the body is not the same as during normal walking. Stopping will give imbalance as well, both due to the inertia of the user and stopping latency.

For Hypothesis Q2, previous works also found differences in presence and cognitive load. We further discuss why we think these will not be a main UX factor that is influenced for our experiment in subsubsection IV-B2 and subsubsection IV-B3. For our measures, many different variables are discussed. These have been organized in Appendix B to Appendix G per-subject: UX questionnaire analysis, VR sickness, time, tracker loss, and order effect, covering: hypotheses, statistical tests, results, and discussion per-subject. The UX questionnaire analysis and VR sickness contribute to our main results. We analyze tracker loss, time, speed, and order effect to discuss whether these extraneous variables influenced our results, and if they impacted user experience. While we did expect tracker loss might reduce the user experience, and a shorter completion time would correlate to a better user experience, we did not expect any of these to significantly differ between our conditions and therefore not to influence our results significantly.

## II. METHOD

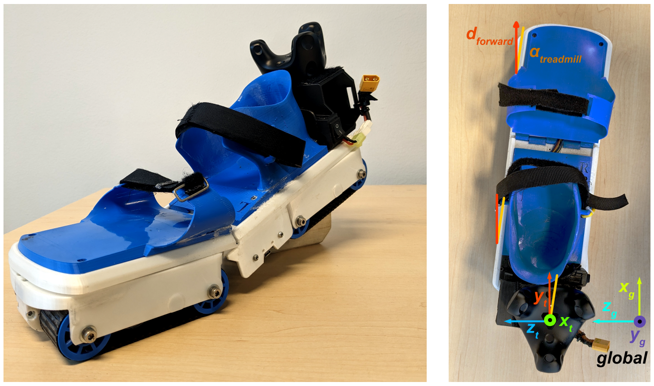
The methods section consists of three main parts. First, the physical device subsection II-A describes the physical VR shoes and how they operate while walking on them. The virtual motion algorithm II-B section introduces the algorithm that maps the positional and rotational tracker data on the shoes, hips and head to the motion in the VR headset. It also describes the four different mappings compared in the user experience test. The user test II-C subsection discusses the experimental procedure, testing environment, task, and outcome metrics.

### A. Locomotion device: VR shoes

The locomotion technique chosen in this work is the Freeaim VR shoes, which are shoes with active (motorized) treadmills underneath (Figure 1a). When both feet are on the ground, the motors are turned off and the shoes cannot move. When one foot is lifted, the other foot will drive backward based on the speed of the lifted foot moving forward. Essentially, this keeps the user from moving forward while walking. The shoe-driving algorithm is proprietary information of Freeaim, and therefore, is not further described in this paper.

The shoe contains two blocks of treadmills, one under the toes, and one under the heels. These can be rotated with respect to the shoe, however, the device version used in this experiment had a fixed treadmill orientation. This orientation was slightly rotated with respect to the foot orientation since this allows the toe to be pointed outwards while walking, see Figure 1b. This also prevents the heels from hitting each other, since the driving motion keeps the feet apart. The shoes were used as a single size for all users. A hinge allows for flexion of the foot, see Figure 1a.

This version of the VR shoes used (from early 2022) did not implement a full positional correction, meaning that the aforementioned algorithm to negate motion was not checked



(a) VR shoe toe joint

(b) $\alpha_{treadmill} = 8°$ to compensate foot rotation outwards.

Fig. 1. The Freeaim VR shoes. This version of the VR shoes had a fixed treadmill angle set at $\alpha_{treadmill} = 8°$ so walking in a straight line, the toes were allowed to point slightly outwards. This is necessary since the shoes otherwise collide easily in a normal gait when driven backward. The toe joint was set at a fixed distance for all users.

against the position in the room. This means that the person could still wander out of the center of the room.

## B. Transfer function algorithm

In this section, we explain the calculation of the locomotion velocity in VR, which consists of the direction and the speed, for which the calculations are explained in the following subsections. The code of this project is freely available at Github [62]. The goal was to develop a mapping algorithm that uses speed and orientation as the main inputs, so it could, in theory, be used with either differentiating position-based trackers as the ones used in this work or with speed or integration of acceleration sensors such as inertial measurement units (IMU's).

The virtual motion seen through the virtual reality headset was generated from the motion and directions of HTC Vive 3.0 trackers [63] positioned on the user's body. One was positioned on the hip, pointing to the front at navel height, and the other two were connected at the back of the VR shoes, pointing in the forward direction, and aligned with the treadmills, see Figure 1b. The head was also tracked using the position and orientation of the virtual reality headset (HTC Vive Pro) [64]. The trackers were tracked using four 2.0 lighthouses/base stations, one in each corner of the room, to prevent tracking loss. The coordinate system referenced in the rest of this section is defined in Figure 1b. Here $x_t, y_t$ and $z_t$ are the local axes of the tracker, and $x_g, y_g$ and $z_g$ are in global (room) coordinate system. The computer used had the following components: GPU 10GB - 3080 (NVIDIA GeForce RTX); CPU 3.70 GHz 12-Core (AMD Ryzen 9 5900X); RAM 32GB.

*1) Locomotion direction calculation / direction condition:* This section discusses the direction used for the VR locomotion velocity vector. Four directions were implemented and compared, shown in Figure 2: i) $\vec{d}_{head}$ (the forward direction of the virtual reality headset), ii) $\vec{d}_{hip}$ (the forward direction of the tracker at navel height), iii) $\vec{d}_{stf}$ (standing foot velocity direction: $-\vec{v}_{standing}$, see subsubsection II-B2), and iv) average feet direction: $\vec{d}_{avg} = \vec{d}_{left} + \vec{d}_{right}$. Due to the tracker often sliding from the hip to navel height in pilots, navel height was determined as the final tracker placement. We also refer to these directions as the direction conditions. These directions were projected on the horizontal plane (so the user would only move horizontally), resulting in $\vec{d}$, which could be $\vec{d}_{head}$, $\vec{d}_{hip}$, $\vec{d}_{stf}$, or $\vec{d}_{avg}$, depending on which direction was enabled. An example of these directions converted to angle outputs, while the user is walking a trial, is given in Figure 18 in Appendix A. The total displacement for a timestep $\delta t$ was then calculated as:

$$displacement = \frac{\vec{d}}{\|\vec{d}\|} * Locomotion\ speed * \delta t, \quad (1)$$

with *Locomotion speed* from Equation 4, discussed in the next subsection.

This displacement from Equation 1 is added up for each frame (integration), resulting in the locomotion position. We use this to calculate the users position, while we keep enabled that if the user moves their head in the room, the optical flow is represented as it naturally would. Therefore if the user wanders out of the center of the room if the shoes do not negate their movement fully, this is still represented in VR. The total sum of the locomotion position and the head position, we call the player position.

*2) Speed calculation:* The length of the locomotion velocity vector is the locomotion speed, for which the calculation is discussed in this section. We give a more in-depth explanation of the algorithm design decisions for each step in Appendix A, including why both the lifted and standing foot speeds are used. In short, this allows for a smoother and more responsive final locomotion speed.

The feet' velocities were obtained by discrete numerical derivation of the tracker positions at 90 Hz in Unity, resulting in $\vec{v}_{lifted\_tracker}$ and $\vec{v}_{standing\_tracker}$. Figure 3 shows the conversion from tracker velocities to the signed speed, $s_{l/r}$, of the left, $l$, and right, $r$, foot, including the following conversions. The standing foot velocity is $\vec{v}_{standing} = \vec{v}_{standing\_tracker}$. Since the user is continuously driven backward by the VR shoe, taking the raw tracker velocity of the lifted foot would be close to zero over time. Therefore, the net lifted foot velocity was obtained by adding the negative $\vec{v}_{standing}$ to the velocity of the lifted foot, resulting in $\vec{v}_{lifted}$ in the reference frame where the user is not stationary. This gives the final foot velocity for each foot used in the algorithm, which can be distinguished by either: $\vec{v}_{lifted}$ / $\vec{v}_{standing}$, or alternatively $\vec{v}_r$ (right foot velocity)/$\vec{v}_l$ (left foot velocity) depending on the conditions of the foot. To obtain $\vec{hv}_{l/r}$, the horizontal velocity of each foot, $\vec{v}_{l/r}$ was projected on the horizontal plane. To calculate $s_{l/r}$, the signed speed for each foot, the magnitude of $\vec{hv}_{l/r}$ was given a positive sign if pointing towards the front of the shoe, and a negative sign

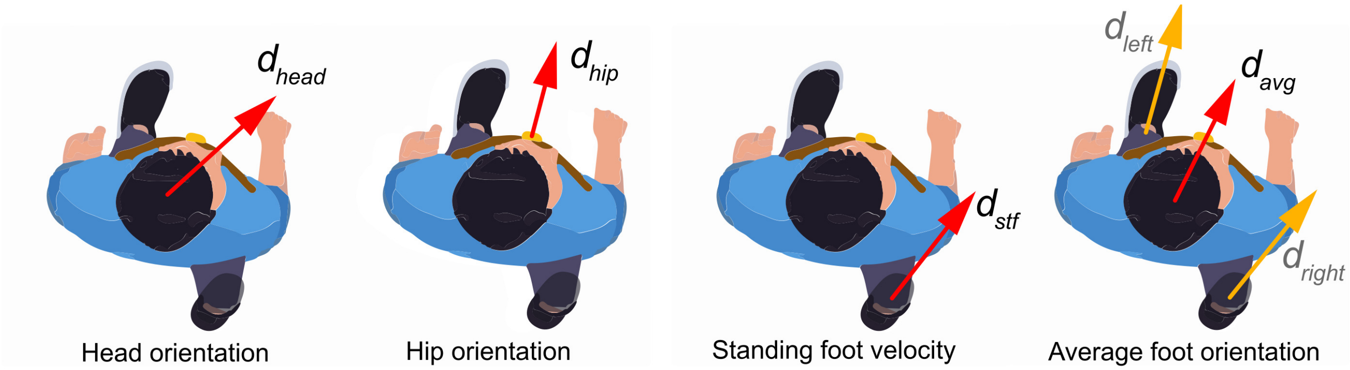


Fig. 2. These are the directions conditions (head, hip, standing foot velocity, or average feet orientation) compared for motion seen through the VR headset. This direction $\vec{d}$ is used as the direction component of the VR locomotion velocity, and is combined with the locomotion speed given in Equation 4 as the length of the locomotion velocity vector.

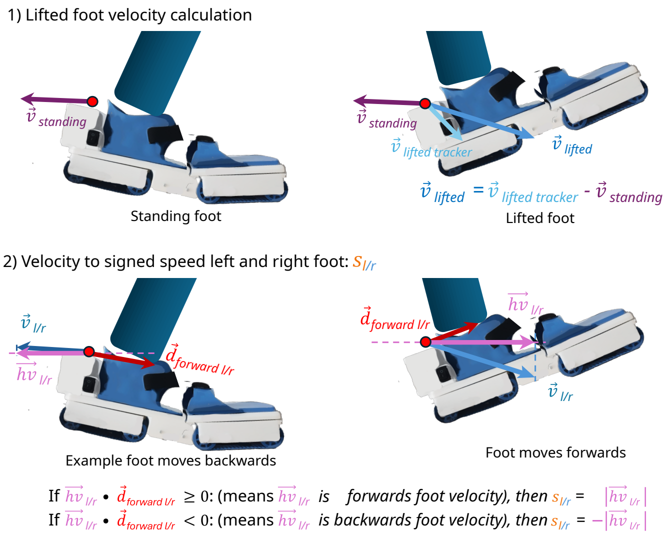


Fig. 3. This figure explains how the velocity from the trackers was processed to the signed speed, $s_{l/r}$, for the left, $l$, and right, $r$, foot. If only the standing foot speed is considered, this gives oscillations in the final locomotion speed. Therefore, to cancel out these oscillations and to respond faster, both the standing and lifted foot velocity, $\vec{v}_{lifted/standing}$, are used to calculate the locomotion speed. $v_{tracker}$ is the tracker velocity in the room/stationary reference frame. In step 1) we convert the lifted foot tracker velocity, $\vec{v}_{lifted\_tracker}$, to the reference frame where the user would not be stationary. For this $-\vec{v}_{standing}$ is added to the $\vec{v}_{lifted\_tracker}$. $v_{standing} = \vec{v}_{standing\_tracker}$ (so kept in the roomscale stationary reference frame), since also converting this to the non-stationary reference frame would result in a standing foot speed of 0. This $\vec{v}_{lifted/standing}$ is the same as $\vec{v}_{l/r}$, left/right foot velocity, depending on the state of that shoe. Step 2) converts $v_{l/r}$ to $s_{l/r}$, the scaled speed. $v_{l/r}$ is projected to the horizontal plane resulting in $\vec{hv}_{l/r}$. $s_{l/r}$ is the signed speed of $\vec{hv}_{l/r}$, positive if $\vec{hv}_{l/r}$ is pointing forwards, and negative if $\vec{hv}_{l/r}$ is pointing backwards. The dot product of $d_{forward\_l/r} = y_{t\_l/r}$ with $\vec{hv}_{l/r}$ is used for determining the forward/backward orientation. The resulting $s_{l/r}$ is plotted in Figure 4.

when pointing towards the back of the shoe:

$$s_{l/r} = \begin{cases} |\vec{hv}_{l/r}| & \text{if } \vec{hv}_{l/r} \cdot \vec{d}_{forward\_l/r} \geq 0 \\ -|\vec{hv}_{l/r}| & \text{otherwise} \end{cases} \tag{2}$$

where $\vec{d}_{forward\_l/r} = \vec{y_t}$ in Figure 1a, and $\cdot$ the dot product. The resulting $s_{l/r}$ is plotted in Figure 4.

To average out noise, see subsection A-E for where we believe this originated from, an exponentially weighted moving average (EWMA) function was applied to $s_l$ and $s_r$ separately, resulting in $EWMA_l$ and $EWMA_r$.

$$EWMA^i_{l/r} = \rho * EWMA^{i-1}_{l/r} + (1 - \rho) * |s^i_{l/r}| \tag{3}$$

where $i$ is the current frame number. The gain $\rho$ was experimentally determined to smooth out the motion as much as possible, without creating a noticeably large response delay in starting or stopping the motion, $\rho = 0.85$, see Figure 15 in Appendix A. As is also discussed in subsection A-F in Appendix A, to calculate the locomotion speed, the $EWMA_{l/r}$ of both feet is combined to cancel out oscillations, and increase responsiveness.

$$Locomotion\ speed = (EWMA^i_l + EWMA^i_r)/4 \tag{4}$$

The factor 4 in Equation 4 is due to a factor 2 to obtain the average, and an additional factor 2 since taking the absolute of a positive and negative speed (one foot moves forwards, the other backward) would give twice the expected walking speed. We tested the resulting speed in subsection A-G in Appendix A. To counter some of these inaccuracies (discussed in Appendix A), this speed was then scaled slightly (1.2 factor) to match the standing foot tracker position to a fixed position on the virtual floor while walking, utilizing pilot tracker recordings of walking on the VR shoes, see subsection A-I on where we believe this mismatch came from. The resulting locomotion speed is plotted with the $s_{l/r}$ in Figure 4.

*3) Determining if a foot is Lifted or Standing:* Determining if a foot was standing was needed for the standing foot direction (to determine which of the two feet would be controlling the direction). Determining if a foot was lifted was needed to obtain the lifted foot velocity, shown in Figure 3. When the VR environment was started, both VR shoes stood on the ground, and a calibration step registered the height of each foot tracker $y_g$, defined in Figure 1a, resulting in: $calib_{yg}$, which remained constant after this initial calibration.

We created two different thresholds to detect if a foot was on the ground. These thresholds and calculations were determined by changing colors of the feet in Unity and seeing if this matched it being lifted or standing, also during walking on

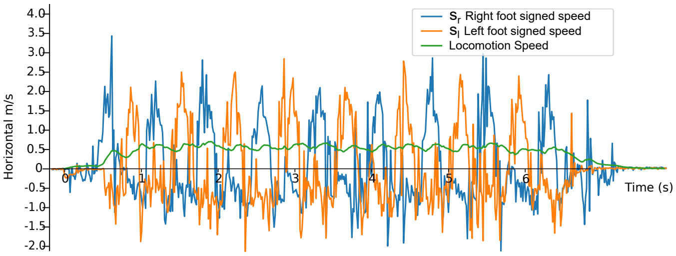


Fig. 4. Horizontal signed speed ($s_{l/r}$) of the feet in the room reference frame, and the resulting virtual locomotion speed from a user walking on the VR shoes. This includes starting and stopping. The negative speeds are due to the shoes driving the user backward while walking and the positive speeds are the forward motion in the swing phase of the user.

the VR shoes, see Appendix A for a further discussion on this. For calculating $\vec{v}_{lifted}$ (see 3), the detection of lifting foot was checked with a stricter threshold $t_{lf} = 0.03\,m$ to prevent adding standing foot speed if the toe of the foot was still on the ground:

$$is\_foot\_lifted = \begin{cases} True & \text{if } yg_{shoe} > calib_{yg} + t_{lf} \\ False & \text{otherwise} \end{cases} \quad (5)$$

To determine whether a foot was standing we used a similar check:

$$is\_foot\_standing = \begin{cases} True & \text{if } yg_{shoe} < calib_{yg} + t_{sf} \\ & \text{and } z_{t\_rotation} > 1^\circ \\ False & \text{otherwise} \end{cases} \quad (6)$$

A more lenient threshold was used since the heel could be slightly lifted in the heel-off phase: $tsf = 0.1\,m$. To prevent false detection in the heel-strike phase when the foot is still lifted, the rotation of the foot was also considered since the toe is pointed upwards just before/during this heel strike. This threshold set for $z_t$ (Figure 1a) was $> 1^\circ$, meaning the toe was pointing down and the heel pointing up, which was the case for the tracker position during standing and heel-off phase.

If both feet were standing (this occurs during the double stance phase in a normal gait), the back foot was taken as the leading foot to determine the direction. The backfoot was determined with:

$$is\_back\_foot = \begin{cases} True & \text{if } (\vec{pg_l} - \vec{pg_r}) \cdot \vec{d}_{forward\_r} < 0 \\ False & \text{otherwise} \end{cases} \quad (7)$$

where $\vec{pg_l}$ and $\vec{pg_r}$ are the global position of the left and right foot respectively, and $\vec{d}_{forward\_r}$ the forward axis of the right shoe.

## C. User test

The anonymous data of the user test is freely available for future research, including the questionnaires and tracking data which can be replayed in Unity. The participants (aside from one, whose data is not shared publicly aside from in this paper) agreed to this. It is available at Github [62].

*1) Participants:* Informed consent was obtained from each subject before the evaluation session from all participants, and we obtained ethical approval from the Human Research Ethics Committee at TU Delft. The experiment was performed with 22 participants. All participants gave permission to use their data in this experiment. The invitation asked for participants weighing a maximum of 80 kg to prevent breaking the VR shoes. During the experiment, participants were weighted, and only accepted with a weight of under 85 kg. Two participants, 62 and 65 years old, were excluded from the experiment since they could not learn to walk on the VR shoes without a walker within 30 minutes. They were replaced by two other participants. The 20 participants (13 male; 6 female; 1 non-binary) between 18 and 47 years old were included in the statistical analysis. One person never used VR, 45% tried it a few times, 25% used VR monthly, 10% used VR weekly, and 15% used VR daily. One participant never used VR. 40% never played 3D video games, 10% played a few times a year, 15% played monthly, 25% played weekly, and 10% played daily. They were also asked for roller skating/skeeler/ice skating experience, and skiing experience. Additionally, body measures were noted: height (164 to 186.5 cm), leg length until the hip rotation point (59 to 99 cm), shoe size (38 to 45.5 EU), foot size (23.5 to 28 cm), and weight (57.6 to 84 kg).

*2) Physical environment:* The test room was a room with 430 x 230 cm of available tracked area the participant could walk in. In the middle over the longest length, there was a safety rail with rollers to attach to the harness that participants wore at all times. The harness rope gave some slack not to feel restricted but would prevent people from falling on their knees if they lost their balance. When walking near the edges of the narrow ends, the harness rope is taut and pulled backward before they can reach a wall. On the long side, the researcher would tell the participant to stop, and then activate a script so the user could walk back to the center of the room without moving in the virtual world or recording tracker and time measurements. Once repositioned, the participant could continue again.

*3) Experimental protocol:* An overview of the experimental protocol can be found in Figure 5. The participants first filled in a consent form, and then filled in a demographic questionnaire on their body weight and sizes mentioned in section II-C1. Once donned a safety harness, they put on the VR shoes and were hooked up to the rail and given a walker for stability to learn the motions, see Figure 6.

The first part of the habituation was done without VR headset. The participants were instructed that two feet on the ground stopped the motion, and the goal was to walk while staying in the same place in the physical room. Since the VR shoes had limited speed, this meant participants often had to adjust their walking speed to the shoes and the first habituation phase was for learning that. When one foot was lifted, the other foot would start to drive. Additionally, they were advised to walk upright since some participants leaned forward. Their goal was to let go of the walker as soon as they felt comfortable. Once they started walking during the habituation phase, the timer recording 'time until no support' was started. The participants practiced walking straight and were asked to let go of the walker if they could. The time until no support was noted at the last time they touched the walker. Once they could walk straight and stop without touching the walker, they were asked if they agreed with the walker to be taken out of reach and then were asked to turn around on the spot slowly while walking clockwise and then a full rotation anticlockwise. They could continue to practice until they felt

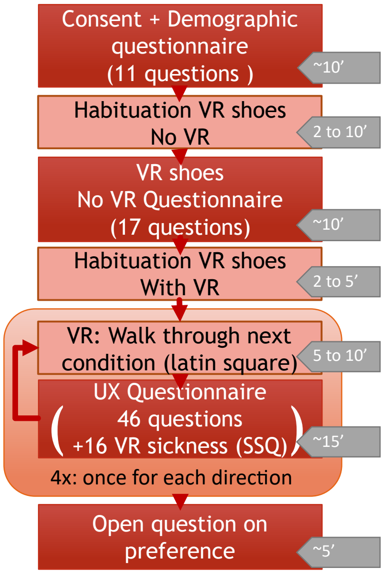


Fig. 5. Overview of the experimental protocol. The total duration of the experiment was 1.5 to 2 hours. The approximate time in minutes of each step is given in the grey boxes, indicating how long the trials lasted, and how long the ‘breaks’ were while filling the questionnaires.

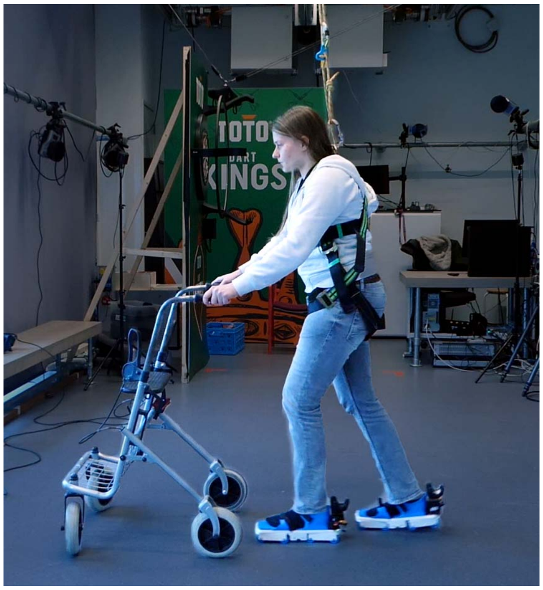


Fig. 6. Habitiuation with the walker. The experiment itself was performed without the walker.

comfortable enough to try walking in VR.

Once they felt confident enough, this time was noted, and they were released from the shoes and ceiling hook. They filled in a questionnaire containing questions on the user experience (UX) of walking on VR shoes (without being in VR). This consisted of the same questions as after the VR experience, as long as the questions made sense. See the overview of UX factors in Table I and the complete questionnaire in Appendix H. Then, they donned the VR shoes, ceiling hook, and VR headset. Now a virtual training environment was started and set to the first direction condition they would perform. The order of the condition was determined by the Latin square, to counterbalance the order of the conditions within a subject. They were given back the walker and again asked to let go as soon as they felt comfortable. The training environment during the VR habituation was an example of following a line on the floor to a circle in front of an artwork, similar to multi-straight line in Figure 7. After standing in the circle and looking at the artwork for a few seconds, the next target (the line/circle to the next artwork) would spawn. This was similar to the first part of the real VR experiment. To practice walking in curves, a curve was displayed on the floor with increasingly sharper turns they could follow for practice. After, they could walk around as long as they wished until they felt comfortable for the real experiment.

Once they felt comfortable enough to walk with the VR shoes in the VR environment, the real experiment was started, with direction conditions (trials) changing one by one following the Latin square order. This experimental phase was performed in the virtual environment depicted in Figure 7. Note that this phase was performed fully without the walker. During each trial (condition), the completion time was recorded between the start and stop, as were the tracker positions and tracker losses in each frame. Tracker losses were detected with the *TrackedDevicePose_t.bPoseIsValid* function from the OpenVR library. The number of tracker losses was filtered in post-processing to the number of tracker losses longer than $\geq 200\,ms$ for each trial. After each trial, participants were released from the ceiling hook and shoes, and filled in a questionnaire (see subsubsection II-C5), after which they performed the next trial/condition. The participants were not told what the difference was between the conditions.

*4) Task and virtual environment:* The user test environment was adapted from [14]. The tasks included in this experiment were selected from their direction control scenario. The multi-straight line and curved walking tasks were used, see Figure 7. Multi-straight line walking let people see the line to follow and circles indicating intermediate goals until the next artwork. Once arrived in the circle of they artwork, they had to wait and look at the artwork for a few seconds, after which the next goal appeared. The curved hallway was a continuous hallway towards the end which required no stops. Cannavò, Calandra, Pratticò, *et al.* [14] also included a backward walking task in the middle, which was ignored, and done forwards instead since the current version of the shoes and virtual algorithm did not support backward walking. A video of the walkthrough in the VR environment on VR shoes can be found here.

*5) Outcome metrics:* Since we are interested in what user experience factors are changing with optical flow, we measured 20 different user experience factors, see Table I, with the questionnaire (Appendix H) adapted from Cannavò, Calandra, Pratticò, *et al.* [14]. This is also used to answer which of our optical flow mappings gives the best user experience. We here explain a few of these terms that could be unfamiliar, if other terms are unclear, we refer to the questionnaire in Appendix H. Appropriateness is about preferring the locomotion technique for this task and being able to achieve what they want. Physical strain similarity asks about the physical strain compared to normal walking. Presence is a common term in VR, and describes the sense of “being in” the virtual environment.

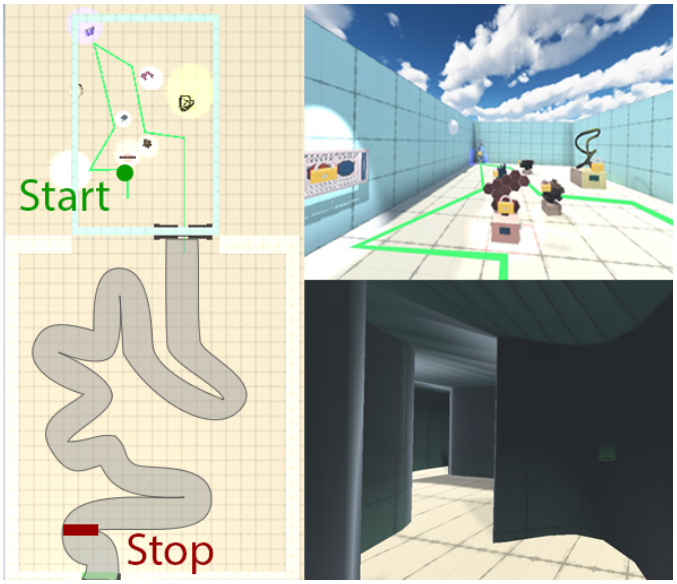


Fig. 7. This figure shows the virtual environment map (left) and first-person screenshots for the multi-straight line and curved walking scenarios.

Here it also contains if the user knew where they were in the environment and if they had a good sense of scale while moving and interacting. Satisfaction asks about how satisfying it is to perform the task, and if the interface behaved as expected. Learnability is how difficult it is to learn the interface, while Intuitiveness is whether the idea of how it works is clear.

Although Cannavò, Calandra, Patticò, *et al.* [14] composed the questionnaire, the original questions are from [56], [65]–[67], which are indicated in Table I as well as the number of questions included for each question group (UX factor). All questions were answered with a Likert scale between 1 to 5, except for VR sickness (SSQ) which had a range of 0 to 3. Three additional questions (see Appendix H) were asked to check how well the speed and direction of the virtual motion algorithm matched the user's perception. The numbers were linked to statements (for example Strongly Agree/Strongly disagree), which is mentioned for each question in Appendix H. Speed was a special case, since 3 was the best result (not to slow/too fast), since 1 was (much too slow/much slower) and 5 (much too fast/much faster).

All question scores were flipped when negative and then scaled to values between 0 and 100, so 100 is best, while the speed questions where the middle value was best (not too slow and not too fast) were processed separately. All individual questions were considered in the statistical model for more statistical power as discussed in subsection II-D, so no averages were taken.

Aside from our score of 0 (maximum VR sickness) 100 (minimum VR sickness), we also provided the SSQ score to compare with previous works. For this, we compute three options: one score where we subtract the baseline SSQ taken before walking in VR and one without the subtraction. Thereby we calculate the score both including "sweating" and without, since we expect this could be due to exercise instead.

At the end of the experiment, three open questions were asked, so participants could mention their observations and their favorite condition in the experiment, see Appendix H. When they arrived at these questions, the researcher prompted them to answer them verbally, asking for more elaboration and clarification before the participants wrote down their observations in the open questions. To calculate the preference score of a condition, it gained one point each time it was considered the favorite condition of a participant. If a participant noted down multiple conditions as favorites, without indicating a clear preference between them, they were each counted partially (for example 0.5 points per condition when two techniques were mentioned) for the preference score.

### D. Statistical analysis

All statistical analyses were performed in Rstudio with R version 4.3.3, using the lme4 and emmeans libraries. The detailed statistical method including relevant R commands, results for all hypotheses, and full discussions are given in Appendix B to Appendix G. The code is also available on Github [62]. Here, a summary is given for the most important tests. The threshold for concluding statistical significance was set at $p < 0.05$ for all comparisons. We organized the following subsections based on the research question it belongs to.

*1) Statistics: Which orientation mapping, choosing from headset, standing foot, average shoes, or hip orientation to VR optical flow gives the best user experience while walking on VR shoes?:* To calculate significant differences between direction conditions in the overall user experience, a linear mixed model (LMM) was used where the direction condition and question group are fixed effects, with the participant ID and specific questions within a question group (UX factor) as random effects:

```
UXscore ~ Direction * QuestionGroup + (1 |
    ParticipantID) + (1 | SubQuestion)
```

The directions were then pairwise compared. All comparisons in this work were adjusted for multiple comparisons using the Bonferroni method within the comparison unless mentioned otherwise. This resulted in a comparison where all UX factors are weighed equally for the overall UX score. Additionally, to see if direction was experienced differently between direction conditions, this question was considered as an additional UX factor in subsubsection II-D2.

*2) Statistics: What UX factors are most influenced by differences in optical flow while walking on VR shoes?:* Here we use the same LMM as in II-D1, but the conditions were pairwise compared for each individual question group, see Appendix B for the R command.

*3) Statistics on extraneous variable: User experience of Speed:* Our hypothesis was that the user experience of speed would not differ between conditions, and would be natural: not too slow and not too fast, since we attempted to map speed 1:1. To see if the user experience would differ between conditions for each speed question, we created a linear model with the interaction of the individual questions and the direction condition. Then we compared the average response of the participants to each speed question between the direction conditions, see section C for the exact R command. The averages of the speed questions were calculated to see if participants felt it was too fast/slow and wanted to go faster/slower overall, while we also looked at the percentages in the number of participants that selected each answer on the Likert scale to see how this was distributed.

*4) Statistics on extraneous variable: Influence of completion time on UX score:* We expected the completion time to not significantly differ between direction conditions, or to be the shortest for the best experienced condition. To compare the completion time between directions, a linear mixed model was made, with the participant id as random effect:

```
completionTime ~ Direction + (1 |
    ParticipantID)
```

This linear mixed model was again used for pairwise comparison on completion time between direction conditions.

We also expected the completion time to be faster for the conditions with the best user experience. To test if a longer completion time indeed correlated with a worse user experience for certain UX factors, a correlation matrix was made between these and (among others) completion time, see Figure 25 in Appendix B.

*5) Statistics on extraneous variable: Influence of tracker losses on our user experience findings:* To test if the tracker losses could have significantly impacted our findings where we compare between direction conditions, the number of tracker losses was pairwise compared between conditions via a linear mixed model, see Appendix F. For this comparison, no multiple comparisons adjustment was applied, since we expected to find insignificance, and adjustment makes it easier to find insignificance. Therefore doing no multiple comparison adjustment is stricter to form our conclusions. To test if tracker losses had a negative impact on any user experience factors, a correlation matrix was made between UX factors and (among others) tracking loss, see Figure 25 in Appendix B.

*6) Statistics on extraneous variable: Influence of order number on the user experience:* This was a within-subject study: participants did multiple direction condition tests in a row. This order was counterbalanced with a latin square (16 orders), repeating the first four orders to balance 20 participants. The order number we refer to in the rest of the document is the number related to the trial performed (1 for the first trial, and 4 for the last trial) within a participant. The order number analyses are discussed in the section G for the overall user experience, VR sickness, completion time, and both individual questions on speed.

## III. RESULTS

The results are organized by each research question, and within the research questions by the hypotheses we test. After the research questions, we show results for some additional extraneous variables (speed, completion time, tracker loss, and order number) that could have impacted our results. The scores were normalized between 0 and 100, so a higher score is a better user experience.

TABLE I
THIS TABLE GIVES AN OVERVIEW OF ALL USER EXPERIENCE FACTORS MEASURED BY THE QUESTIONNAIRE. THIS QUESTIONNAIRE WAS COMPOSED BY [14]. THE ORIGINAL QUESTIONNAIRES CORRESPONDING TO THE QUESTIONS ARE NOTED BETWEEN THE BRACKETS AS VRUSE [65], NILSSON [56], AND ISO [66]. THE NUMBERS INDICATE THE NUMBER OF QUESTIONS FOR EACH FACTOR.

| | |
|---|---|
| Ease of Use (VRUse, 9) | Naturalness (Nilsson, 1) |
| Perceived Errors (VRUse, 5) | Physical Strain Similarity (Nilsson, 1) |
| Appropriateness (VRUse, 6) | Perceived Physical Effort (ISO, 1) |
| Input Sensitivity (VRUse, 1) | Mental Effort (ISO, 1) |
| Input Responsiveness (VRUse, 2) | Satisfaction (VRUse, 2) |
| Acclimatisation (Nilsson, 1) | Self-motion Compellingness (Nilsson, 1) |
| System Usability (VRUse, 3) | Presence (VRUse, 4) |
| Control (VRUse, 1) | Enjoyability (VRUse, 1) |
| Learnability (VRUse, 1) | Comfort (VRUse, 1) |
| Intuitiveness (VRUse, 3) | VR sickness (SSQ, 16) |

### A. Results: Direction mapping with the best user experience

We expected that the head-oriented walking would give the worst user experience, and standing foot velocity oriented walking would give the best user experience.

When considering all user experience factors (Table I) equally, the conditions of average feet, hip, and standing foot velocity yielded significantly higher scores than head-oriented walking ($p < 0.001$), see Table IV in Appendix B. None of the other conditions were significantly different in pairwise comparisons. For the preference score 6.5 participants indicated average feet as their favorite, followed by hip (5.5), standing foot (3), and head (1). The remaining 4/20 participants did not fill the question in (2), did not notice differences (1), or gave a contradicting answer (1): they mentioned that the third condition (hip) was best, but in the open question after that third condition itself, they mentioned: “This time, turning was very hard. I didn’t feel like I was in control at all”.

### B. Results: User experience factors most impacted by optical flow differences

Each of the following subsections discusses results related to the hypotheses of the question: *What user experience factors are most influenced by differences in optical flow while walking on VR shoes?*

*1) Hypothesis: There are multiple UX factors that are significantly different between direction conditions:* A bar plot with p-values < 0.05 for the UX factor comparisons is shown in Figure 8 sorted based on p-values significance, a bar plot with the largest estimates (similar to differences between means) is shown in Figure 9, sorted based on the largest estimates, and a matrix of means Figure 24 is shown in Appendix B. The pairwise p-value and estimates ($\beta$, similar to the mean difference between conditions) table can be found in Appendix I.

First, we discuss the results in Figure 8. The Ease of Use score for the head direction was significantly lower (harder to use) than the hip ($\beta$ = 9.8, $p$ = 0.001), standing foot ($\beta$ = 9.3, $p$ = 0.003), and average feet ($\beta$ = 8.8, $p$ = 0.006) conditions. The Input Responsiveness was significantly worse for the head compared to the hip ($\beta$ = 15.6, $p$ = 0.011) and standing foot ($\beta$ = 15.6, $p$ = 0.011), and while the head also yielded lower scores compared to the average feet condition, the difference

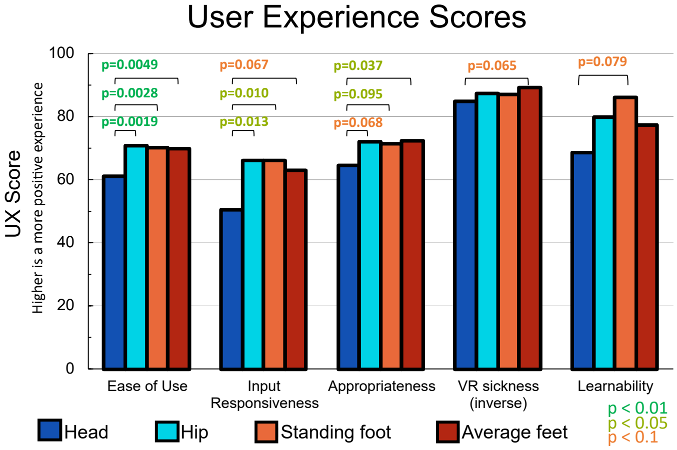


Fig. 8. User experience scores are shown for each direction condition with p-values indicating pairwise comparison scores, for UX factors when any pairwise comparison $p < 0.1$ was found. The colored bars are the four direction conditions we compare. The scores are averages from the questions of the corresponding group and are flipped and scaled to match the 0 to 100 scores, with 100 being the best rated (for example the least VR sickness, or the easiest to use). The head condition was significantly worse compared to all other conditions in Ease of Use. For Input Responsiveness, the head was significantly worse compared to hip and standing foot velocity direction. While for Appropriateness the head is significantly worse compared to standing foot and average feet. Orange values ($p < 0.1$) are not considered significant.

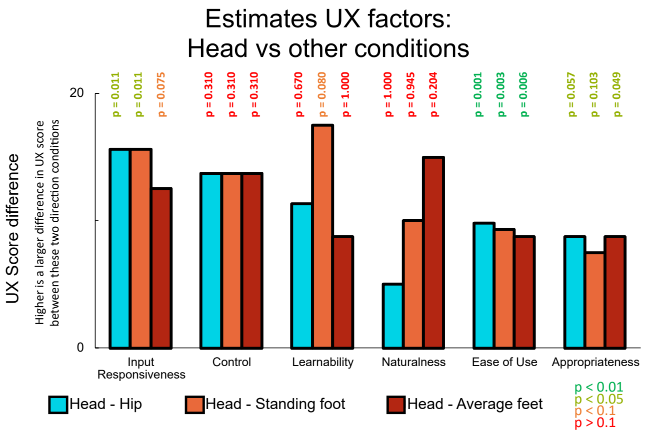


Fig. 9. This figure shows the estimates (similar to differences between means) of the head and each other condition, sorted by their average estimate combining each comparison (head-hip, head-standing foot, head-average feet). The user experience factors displayed were selected by having an estimate larger than 10, or being significant ($p$ < *0.05*). Note that while some have larger mean differences, these differences were not significant while smaller differences were.

was not significant ($\beta = 12.5$, $p = 0.075$). In Appropriateness, the head condition was again rated worse compared to average feet ($\beta = 8.75$, $p = 0.046$), and also to hip ($\beta = 7.5$, $p = 0.057$) and standing foot ($\beta = 6.880$, $p = 0.100$) although not significant. The head condition was scored lower than the standing foot condition in terms of Learnability, meaning head was harder to learn than standing foot, but the difference was not significant ($\beta = 17.5$, $p = 0.080$), with the other Learnability comparisons giving no significant differences ($p > 0.671$).

Therefore there were three UX factors significantly different between conditions: Ease of Use, Input Responsiveness, and Appropriateness. None of the other tested 17 UX factors demonstrated significant results, aside from the few $p \leq 0.1$ mentioned, all being $p > 0.150$. This includes no significant difference between direction conditions for the question explicitly asking if participants felt they were moving in the direction they wanted to go ($\beta \leq 11.3$, $p > 0.67$). However, direction/turning differences between conditions were noticed in the open questions by 8/20 participants.

Note that while the UX factors above were the (near) significant differences, there were UX factors with larger estimates, see Figure 9, where they are sorted based on their averaged estimates between each comparison. The three largest individual estimates between head versus another condition (hip/standing foot/average feet) being Learnability (to standing foot $\beta = 17.5$, $p = 0.080$), Input Responsiveness (to standing foot $\beta = 15.6$, $p = 0.010$), and Naturalness (to average shoes $\beta = 15.0$, $p = 0.200$).

*2) Hypothesis: VR sickness is the most influenced UX factor by differences in optical flow:* There were no significant differences in VR sickness between conditions ($p > 0.092$). The head condition received lower scores for VR sickness compared to the average feet condition, meaning more VR sickness occurred for the head than average feet, although not with a significant difference ($\beta = 4.3$, $p = 0.092$). The rest of the conditions showed no significant difference in VR sickness ($p = 1.000$). The overall SSQ score without sweating varied between -18 (meaning VR sickness got better compared to before walking in VR), the average score was 14, and a single participant scored high on VR sickness (119), followed by other participants with an SSQ sickness of a maximum 60. The SSQ score including sweating was similar and varied between -22, an average of 16, and a maximum of 123. Without subtraction of the baseline and including sweating, SSQ was between 0, an average of 28, and a maximum of 123.

### C. Results: User experience of the VR shoes

In this part, we show the results relating to *What are the current user experience limitations of VR shoes?* For our hypothesis, we expected that users would have to learn to walk on the VR shoes, and might have problems with balance at the start since there are a few differences in motion between walking on the VR shoes and walking normally. Additionally, we expected difficulty during stopping.

*1) Results: Overall VR shoes user experience with VR:* For the overall walking on VR shoes we use the results presented in subsubsection III-B1 for the best overall direction conditions from subsection III-A, namely hip, standing foot, average feet. The direction scored 80/100 on average for these conditions in how well people found it matched the direction they wanted to go after walking in VR. One participant also mentioned that it felt like the standing foot condition had an offset with respect to their desired location.

*2) Results: Habituation - observations on learning to walk on VR shoes:* The time until no walker support required without VR was at a minimum of 20 seconds, on average

1 minute and 52 seconds, and at a maximum of five and a half minutes for the younger participants (max 47 years old). The full training without VR took a minimum of 72 seconds, on average four and a half minutes, and a maximum of 10 minutes. However, two participants 62 and 65 years old did not manage to complete the training even with 30 minutes of practice. Therefore they were excluded from the experiment, and replaced by two other participants.

When learning to walk on the shoes with no VR, many users needed to actively practice adjusting to the speed of the VR shoes, since they kept wandering from the center of the room forwards or backwards. Some users mentioned the time they took to swing their leg forward was shorter than the time the VR shoe took to drive their other foot back. Participants also mentioned they had to take shorter steps and had to plant their foot straight down instead of their natural lifted foot swing. On average users felt the shoes were moving at a comfortable speed (3.2, on a scale of 'wanted to go much slower' (1), and (5) 'much faster'). However, 8/20 participants said they wanted to go faster, including one much faster, while 7/20 participants wanted to go slower. In the open questions, one participant mentioned they wanted to go slower in the beginning, but got used to it.

Participants also mentioned having trouble with stopping: they felt like they were going backward, and that the shoes only stopped after a "long stop reaction time". For turning, a participant mentioned that some caution was required in foot placement while turning so one shoe does not back into the other, and for multiple participants, we saw their heels colliding while practicing turning. A few participants had to adjust their foot abduction so their heels would not bump into each other while walking straight. Participants also mentioned similarities to ice skating and skiing in the open questions. Two participants described the experience as a bit like ice skating, another as walking like a robot, while another mentioned it felt natural once you get used to the shoes but with some skiing elements to it.

### D. Results on extraneous variables

Several additional variables were investigated (speed, completion time, tracker loss, and order effect), to see if these could have impacted our results. Here we show the results related to each of these variables.

### E. Results on extraneous variables: How did users experience the Virtual Speed, and did this influence our results?

We expected the user experience of virtual walking speeds to be not significantly different between conditions since they were calculated with the same variable. This was indeed the case, see Table V in section C.

Thereby, we expected the speed to be considered natural (not too fast or too slow). The optimal score for speed was 3, with higher (maximum 5) being too fast or and lower than 3 (minimum 1) too slow. The virtual speed was perceived as slightly slower than normal walking for all conditions (average 2.66 over all conditions) with no significant difference between the conditions ($p > 0.53$). Combining all conditions, 36% of participants found the virtual speed felt natural compared to normal walking, 21% felt it was faster, while 42% felt it slower of which 13% much too slow. Users wanted to go slightly faster virtually for all conditions (3.47 average), with no significant difference between the conditions ($p > 0.95$), see Appendix C. Of all participants, 30% were happy with the virtual speed, 54% wanted to go faster in VR, of which 9% wanted to go much faster. There were some participants (16%) who wanted to go slower, and no one wanted to go much faster.

Some participants thought the main difference between conditions was speed and were comparing them as such in the open questions. 7/20 participants mentioned in the open questions that they thought that speed was the main difference between at least two conditions.

We expected users to get used to the virtual speed and want to go faster for a higher order number. This was not the case ($p > 0.1$). However, since this is an analysis of a single question for which we have seen our sample size is not big enough to obtain significance, it could be this was affected but is not visible due to the statistical power.

### F. Results on extraneous variables: Did the Completion Time influence our results or the UX?

We expected completion time either to have no significant differences, or to be faster for the condition with the best user experience. We also expected that a faster completion time would positively correlate with other user experience factors. The trials did not significantly differ between conditions ($p > 0.8$), also see Table VII in Appendix E, and lasted between two and a half minutes, and seven minutes (four minutes on average). A faster trial completion time correlated negatively with Appropriateness ($p < 0.01$), Direction ($p < 0.05$), and Enjoyability ($p < 0.05$), see Figure 25 in Appendix B.

### G. Results on extraneous variables: Did the Tracker Loss influence our results or the UX?

We expected tracker loss to have no significant difference between conditions, and that a more tracker losses would increase VR sickness and correlate negatively with many user experience factors. As visible in Table VIII in Appendix F, the tracker loss was not significantly different between conditions ($p < 0.59$). Tracker loss or resulting behavior (see subsection IV-G) was mentioned by 9/20 participants in the open questions. As visible in Figure 25, more tracker losses only negatively correlated with the questions related to Control ($p < 0.01$) and Self-motion Compellingness ($p < 0.01$). There was no significant correlation with any of the other UX factors.

### H. Results on extraneous variables: Did the Order Effects influence the UX?

The order effect results are all calculated and plotted in Appendix G. We investigated the order effect for the overall user experience, VR sickness, completion time, and both individual questions on speed.

## IV. Discussion

In this work, we compared the effect of multiple optical flow direction algorithms on user experience while walking on VR shoes. We compared four different optical flow algorithms, varying the direction based on the orientation of body-mounted trackers, namely the head, hip, standing foot velocity, and average feet. The user experience was measured with a questionnaire (Appendix H) measuring 20 user experience factors, and multiple open questions on their experience and preference. With these results we evaluate three things: which of the tested optical flow direction mappings gives the best user experience? What user experience factors are most influenced by differences in optical flow? And, what are the current limitations of VR shoes? After discussing the answers to these questions, we also investigate some extraneous variables that could have impacted our results for these questions.

### A. Which Optical Flow Direction mapping resulted in the best User Experience?

Direction optical flow movement based on the head orientation was experienced worse compared to the other conditions (hip orientation, standing foot velocity direction, average shoe orientation). This matched our expectations, since participants could not look in a separate direction from their current movement direction. In normal walking, people first move their eyes towards the target, then the head orientation follows anticipating the future walking direction by 200 to 600ms before the body turns [54], however, with the virtual head orientation, this is not possible, since this would cause the user to immediately start moving in the direction they look at: head orientation and movement direction is not decoupled.

However, in previous walk-in-place (WIP) studies on direction, the head often was found to give a better user experience compared to the torso [15], [59], and average hands direction [15]. There are two main differences between their and our methods. In their algorithm, the tracker was mounted on the torso, while in our experiment it was mounted on the hip (navel height). It has been suggested that torso-mounted tracking causes more unreliable rotations than hip/foot directed steering [22], [60]. Thereby, their walking movement was with walk-in-place, while ours was with full gait. Therefore their motion is more artifical, which might make users act differently and less intuitively while walking-in-place compared to walking on the VR shoes, which could include a different orientation of their body.

The other conditions (hip, standing foot velocity, average feet) did not significantly differ from each other. Our reasoning for this is further discussed in future work: subsection IV-L

### B. Which User Experience Factors were most Influenced by Optical Flow differences?

We investigated the impact of the direction on different UX factors. There are two considerations here: which were significantly different for some conditions, and the estimate (similar to the difference between means) of the values themselves.

We look at the difference between the head and the other three conditions, since here it seems there was a significant difference in user experience between these optical flows. The user experience factors differing most between the head and the other three direction conditions were Ease of Use ($0.001 < p < 0.006$), Input Responsiveness ($0.011 < p < 0.075$), Appropriateness ($0.046 < p < 0.1$), followed by some not fully significant: VR sickness ($0.065 < p \leq 1.000$) and Learnability ($0.080 < p \leq 1.000$). However, this does not mean these were the most influenced UX factors. Looking at the estimates (similar to the difference between means), we can see for example Ease of Use ($8.8 < \beta < 9.8$) and Appropriateness ($6.88 < \beta < 7.5$) had a smaller estimate between direction conditions than Control ($13.7 < \beta < 13.7$), Learnability ($8.75 < \beta < 17.5$), Naturalness ($5 < \beta < 15$), and Direction ($1.25 < \beta < 11.3$), but these were not significantly different between conditions, while the others were. We believe this is due to the number of questions asked: Ease of Use, Appropriateness, and Input Responsiveness had a larger number of questions (7, 6, and 2 questions respectively), versus only a single question for for example Naturalness. Due to the extra data points these give, there is less chance a difference is due to noise, and therefore the p-values will be lower. Therefore, aside from naming which user experience factors significantly differed between user experience factors (Ease of Use, Input Responsiveness, and Appropriateness), and naming the largest estimates although not always significant (Input Responsiveness, Control, Learnability, Naturalness, Ease of Use, Appropriateness, and Direction), no clear conclusion can be drawn on the user experience factors that were most impacted by the optical flow (varying direction conditions).

*1) VR sickness: Expectation versus reality:* Due to previous works on optical flow, we expected VR sickness to be influenced most [20], [21], especially since a conflict between optical flow and visual input is thought to cause VR sickness [18]. However, as we have seen, VR sickness was not fully significantly different between conditions ($\beta = 4.3$, $p = 0.092$ between head and hip, and $\beta \leq 2.21$, $p = 1.000$ for the other tests). Thereby, we expected VR sickness to increase with the order number (it being the nth trial for that user), since the total VR exposure time becomes longer with each trial, but this was also not significantly different. In fact Figure 27 shows that for most participants the VR sickness flatlined or even improved (although not significantly). However, especially for participant 14, this was not the case, which we at the bottom of this section. The VR sickness was also relatively low for a VR locomotion experience. For example, Lochner and Gain [68] had a two to three minute duration experiment, and measured SSQ scores of 33 for normal walking, 32 for armswing, 49 for joystick locomotion, and 24 for teleportation. Even without substracting the baseline, our SSQ came to an average of 28, while our experiment was four minutes on average. Of course, they had a different task and environment than we did, which could have caused these differences.

It could be that since the VR sickness score was relatively low, differences were so minimal they were barely measurable. We think the low VR sickness scores could be because of four factors. First, it could be due to the short time in VR of only four minutes per trial on average, with a break of around fifteen minutes filling in the questionnaire between each trial.

In previous work, VR sickness seems to increase after the first ten to twenty minutes in VR [69], however, Lochner and Gain [68] showed that for VR locomotion, even after three minutes, there is already some presence of VR sickness.

Second, physically walking [50], or stimulating the vestibular system with vibration on footfalls [70] seems to reduce VR sickness. Slidemills also fulfill these criteria but do not reduce VR sickness [6]–[8]. However, we believe this is due to the waist restraint, preventing vertical motion of the upper body and head, including the vestibular system. Since the VR shoes have no such restraint, and foot impact traveling up the spine that can cause similar vestibular vibrations, we expect that walking on VR shoes in result in relatively low VR sickness compared to other techniques such as joystick locomotion, or slidemills. This is supported, as we can also see from the results by Lochner and Gain [68] for normal walking (SSQ 33) compared to joystick (SSQ 40).

Third, our participants were relatively experienced with VR, with only one person never having experienced VR, and 50% of the participants using VR at least monthly. VR sickness reduces with repeated exposure [71], [72], which means people more experienced with VR likely experience less VR sickness. One participant (14), in particular, got very sick (119 to 123 SSQ depending on baseline/no baseline/no sweating), while the rest was below 60 in all cases. This participant was the same participant who had not used VR before. Therefore, for VR sickness evaluation, this should be tested with a group with less VR experience, more similar to the general population. Finally, an additional factor that might have prevented finding significance, are the tracker losses (see subsection IV-G). Some participants mentioned that tracker losses caused disorientation, which is one of the symptoms of VR sickness. Therefore, while tracker losses varied, it provided noise in the VR sickness score. Therefore, the VR sickness with already low scores and this additional noise could have caused more difficulty in finding significance.

*2) Mental effort: Expectation versus reality:* Mental Effort was not expected to have a significant difference between optical flow conditions, even though cognitive load was previously found to be impacted by optical flow [21]. We think there are two main reasons for this. First, the experiment by Bruder, Lubos, and Steinicke [21] was performed with a specific memory task. Participants had to walk behind a sign that displayed a sequence of letters over time, and had to press a button when that letter was the same as the letter two iterations before it. Since this task would have distracted from the complete walking experience, we decided to ask this by a question already present in the questionnaire instead. However, without a cognitive challenge, an increase or decrease in mental effort might not have been noticed as much by the participants. A second reason was that there was only one question on Mental Effort, and as we discussed before, this did not result in enough data points to result in statistical significance with our sample size. However, the estimate was also not particularly large ($3.8 < \beta < 7.4$). Therefore we think the former argument is more likely.

*3) Presence: Expectation versus reality:* Presence was previously also found to be broken with mismatching optical flow [16]. Schmitz, Hildebrandt, Valdez, *et al.* [16] varied the rotational gain, meaning scaling the rotation to rotate faster or slower. They found the threshold when presence breaks is much more forgiving (gain under 0.45 or above 1.85) before it breaks presence on 50% of the users, compared to the detection threshold where 50% of users notice the optical flow is manipulated (under 0.8 or above 1.49) [53]. We believe the differences in optical flow were around or slightly under the detection threshold since only 7/20 participants noticed this as a difference between conditions. Therefore it seems logical that the presence was not different between conditions, since the difference in optical flow was not near the level of mismatch that presence would break.

*4) Concluding words on the most influenced user experience factors by optical flow:* We did not find what we expected: our expectation was a high influence of differences in optical flow on VR sickness (and perhaps Mental Effort and Presence) based on previous works, but we did not find significant or large differences for any of these. However, we did find optical flow influences on many user experience factors this has not been researched for: at least Ease of Use, Input Responsiveness, and Appropriateness. Thereby it seems some factors (Control, Learnability, Naturalness, Appropriateness, and Direction) might be influenced even more, although these differences were not significant, which we believe is because of a lack of statistical power in the smaller number of questions in these question groups.

### C. Current User Experience Limitations of the VR shoes

To analyze the current user experience of walking on VR shoes, we are interested in the results from the best direction condition since it is the most likely to be used in the future. Therefore we look at the three lower-body conditions averaged (excluding head), since the head was found significantly worse in user experience than all lower-body conditions. No particular UX factor was outright negative (score < 45). However, Comfort (49), Physical Strain Similarity (51), Perceived Errors (54), Acclimatisation (57), and Naturalness (58), have the lowest scores. This shows that the motion of walking on VR shoes in combination with VR does not completely feel like normal walking. Direction scored an 80/100 for how well people feel it matches after walking in VR, so although it is good, it is not perfect yet. Currently, the VR shoes' movement while driving the foot back has a fixed angle with respect to the foot orientation. During normal walking in a curve, the standing foot can travel in an arc with respect to the body. Allowing the VR shoes to move the foot in different directions independent of foot orientation could improve this if properly tuned so the movement is similar to foot movement directions in overground walking. Additionally, some users had issues with their heels colliding even while walking straight. The angle of the foot with respect to straight movement direction of the treadmill, $a_{treadmill}$, did not seem to be enough for some participants. Since only some users struggled with this, we believe this should be adjustable depending on the user's preferred abduction while walking straight.

The shoes had a limited speed which seemed to additionally depend on the weight of the user and the battery level. Thereby, since no roomscale positional correction was implemented in this prototype of the VR shoes, some users had to walk slower or faster. Adjusting their gait to this faster or slower pace was mentioned as more tiring, which is logical since each person tunes their natural gait to be the most energy efficient for their body [73]. While users were learning to walk on the shoes, an additional issue was that some users leaned forward too much while trying to balance, especially while trying to stop. This caused the heels to not fully reach the floor, causing further forward imbalance as the shoes were not stopping when the user wanted them to. Advising people to stand upright and lean on their heels often helped. To prevent this in future models, we think that the sensor that detects if the VR shoe is on the ground should be placed under the toe part instead of the part hinging with the heel of the VR shoe.

In the VR shoes training without VR, it was often seen that participants had the opposite difficulty keeping balance on stopping. When the shoes suddenly stopped, the momentum of the participant's body kept going, causing the body to go further backward before the user corrected this. After getting used to it, participants leaned slightly forward prior to stopping to compensate for this. A more gradual stop could prevent this (similar to Ekto VR boots[42]), but would cost more room space.

After learning to stop, the next step was turning, which participants also found difficult to learn. Participants mentioned similarities to ice skating and skiing. As mentioned above, the standing foot would drive in a fixed direction, while people walking in a curve on the normal floor would gradually curve their foot trajectory with respect to their body. Since the feet are similarly restricted in foot movement direction to foot orientation during ice skating and skiing, it could be that this comparison is due to this fixed rotation. In the future, this standing foot reorientation motion might be simulated by prediction algorithms, combined with driving the threads individually to make the shoes drive in a curve. One participant mentioned that it felt like the standing foot direction condition had an offset with respect to their desired location. This virtual motion might also improve with this VR shoe improvement.

These differences in gait were learnable for all participants between 18 and 47 years old within five and a half minutes of walking straight, and at maximum 10 minutes until they practiced turning and felt ready for VR. It probably also played a role that most participants were technologically experienced, and more than 75% were accustomed to balancing sports such as roller skating, ice skating, and skiing. While this group could learn the VR shoes in minutes, two elderly participants (62 and 65 years old) were not able to learn to walk on the VR shoes without the walker, even with half an hour of practice. Therefore, in its current form, this is not suitable for the elderly or rehabilitation. Additionally, this was only tested with a weight of up to 85 kg to prevent the shoes from breaking. With future versions of more sturdy shoes, participants with weights above 85 kg should be tested.

The version of Freeaim VR shoes used in this environment could not yet correct the position of the user with respect to the room scale. Participants were secured with a safety harness to a railing that allowed movement throughout the room but prevented a collision with the walls or falling. However, this resulted in participants often wandering from the center of the room, and then being restricted by their safety harness. Aside from being pulled off balance, this sometimes resulted in being lifted from their heels by the safety harness, resulting in runaway shoes. Therefore this heel lifting problem should be considered in any safety rig and future VR shoe designs.

Additionally, a single user mentioned the noise of the VR shoes actively preventing them from feeling present and walking in the virtual environment and another (>60 years old) that it was sometimes preventing them from understanding the instructions in training.

An additional thing that differs from normal walking is that the VR shoes' speed seemed to gradually increase after heel strike instead of giving a constant speed for the standing foot, see Figure 23 in Appendix A, which happens less in for example a high torque standard treadmill.

### D. Extraneous variables

Speed, tracking loss, and completion time are all additional variables that could have impacted our results. In this section, we discuss whether these variables affected the outcome.

### E. How did users experience the Virtual Speed, and did this influence our results?

There are a few measures we consider here: the subjective experience of speed filled in the questionnaire, the completion time, and the observations on the speed of the physical VR shoes. We did not expect any of these speed variables to differ between conditions, and we expected the speed to feel natural since we attempted to make the speed algorithm 1:1. The subjective speed was not significantly different between conditions ($p > 0.53$) so this did not significantly influence the answers to our results where we compared between optical flow directions. However, we do find some interesting observations.

Participants filled in the questionnaire that they felt they were moving slower in VR compared to natural walking, while it was matched 1:1 as well as possible, which should feel natural with the field of view (110°diagonal) used [52]. It could be that since the actual VR shoes were often considered slow, and users had to adjust their speed, they considered the virtual speed "slower compared to natural walking". Some participants mentioned they found the speed in VR similar to what they were walking, but since they were limited by the VR shoe speed, it cost more energy, also see subsection IV-C. Another participant mentioned it did feel like a natural speed, but the distances in VR felt longer and more tiring. One participant mentioned: the speed for the museum part was good, but when walking the long curved hallway it felt too slow. We think this was probably due to walking a long boring path, and possibly a lack of optical flow of the grey walls. However, another participant contradicted this: they felt the virtual environment did not result in the same speed as they were walking, and VR should move faster to feel like it matched. Some said they felt a lack of control over speed, which was probably due to tracker losses since this would cause them to suddenly shoot forward after recovering the tracker, or moving while they did not intend to, see subsection IV-G.

However, we discuss several reasons why participants could have experienced the virtual speed as slower in subsection A-J in Appendix A due to limitations in the speed algorithm, and potential solutions to improve it in subsection A-K. Additionally, we discuss how we could allow for maneuvering (small correction steps), and how to solve what some users mentioned: that they kept moving after stopping.

Some participants thought the main difference between conditions was speed and were comparing them as such in the open questions. However, the completion time, see subsection IV-F, as well as the subjective scores on speed were not significantly different between conditions. Therefore, we believe this variance might have been caused by tracker loss, see subsection IV-G, or the order effect, see subsection IV-H.

### *F. Did differences in completion time influence our results or the user experience?*

We expected completion time to be similar, or perhaps better for a condition with a better user experience. The completion time was not significantly different between the conditions ($p > 0.8$). This means participants were not able to complete the trial faster in one condition than another. Therefore, when participants perceived differences in speed within the questionnaire, this was either averaged out between the conditions, or purely subjective, and therefore did not influence our results.

We expected the participants to be able to complete the trials faster as they gained more experience. This was indeed the case, but only significant for the first to the second trial, indicating their learning curve was mostly leveled out after training one full run in VR.

We also expected that a faster completion time would correlate positively with a better user experience since users were expected to have fewer problems walking when they were able to complete the course faster. Interestingly, we found no significant positive correlations, but we did find that a faster completion time had a significant negative impact on some UX factors: Appropriateness, Enjoyability, and Direction. This could be explained by the participants attempting to rush through the experiment, and in the process having trouble with the VR shoes and keeping balance, or running into other control issues due to this. Especially Appropriateness is about preferring the locomotion for this task and being able to achieve what they want. From this, it is unclear whether being able to achieve the direction they desired is about the virtual direction condition, or having difficulty turning corners while walking on the VR shoes themselves. One user mentioned in the open questions that they slipped away more often during later conditions and had to constantly try to keep balance. However, another participant mentioned their feeling of presence improved growing more comfortable with the shoes and being able to look around more in the virtual environment. Therefore this might depend on the need to rush some participants felt.

### *G. Did tracker loss influence our results or the user experience?*

The result of a tracking loss was that the participant would not move during the tracker loss, but would suddenly shoot forwards with the missed speed between the previous and new tracking position directly after tracking recovery. Tracker losses were mentioned to cause disorientation, loss of balance, and involuntary VR movements in the open questions. Although the tracker losses were often mentioned in the open questions of the questionnaire, they seemed to have had little impact on the user experience scores. First of all, the number of tracker losses was not significantly different between conditions, so it is not expected to have a noticeable impact on results comparing between the conditions. When looking at the correlation with user experience factors combining data from all conditions, only Control and Self-motion Compellingness seemed to have been impacted. Interestingly Perceived Errors, speed, and other seemingly related questions groups did not seem to be impacted. Since for our main research questions we were not specifically interested in the user experience of tracker loss, when a participant asked or mentioned (or behavior off) tracker loss, and if they should include the tracker loss in Likert scale questions on UX factors while filling in the questionnaire, participants were told to try and ignore the instances of tracker loss. Since it does not differ between conditions and does not correlate with most of the UX data but was often mentioned in open questions, this request seems to have reduced tracker loss impact in the questionnaire results.

It seems the tracker losses were due to a combination of multiple factors. One of the lighthouses would sometimes suddenly move to a different place virtually than physically. If this lighthouse was then taken as a reference point by a tracker, the tracker would shift with that distance. Additionally, the right foot tracker was smashed during the pilot by a stumbling user. While it seemed fine, analyzing tracker loss after the experiment showed that mainly the right (smashed) foot lost track often, and may have been internally damaged. Additionally, walking is a quick movement with occlusion behind the legs, which may cause losses even if the tracking system works. In the future, the “shoot forward” effect on tracker loss can be further reduced by programming the speed algorithm to ignore data briefly on tracker loss reconnection, but this does still cause stops on tracker losses. Additionally, another tracking method, such as an IMU not influenced by visual occlusion, could be used to obtain the speeds, although integration errors may build up over time and should be evaluated for this use.

### *H. Did order effects influence our results or the user experience?*

We expected the user experience to be better with a higher order number. Indeed, order numbers 2, 3, and 4 all have a better user experience than trial number 1. Therefore we expect users were still learning to walk on the shoe, which negatively impacted their user experience. This is especially visible since Perceived Errors was an individual UX factor that was significantly different for the first trial compared to all others ($p < 0.003$). We expect this did not influence our experiment however, since we used a Latin square (with participant amount a multiple of 4, for 4 conditions), meaning it should be fully counterbalanced.

The order effect on VR sickness was discussed in subsubsection IV-B1, the order effect on the completion time was discussed in subsection IV-F, and the order effect on virtual speed was discussed in subsection IV-E.

### *I. Concluding words on extraneous variables*

All extraneous variables (virtual speed, completion time, tracker loss, and order effects), had varying influence on the user experience but were not significantly different between the direction conditions. Therefore, we believe they did not impact our results comparing these for the main research questions.

### *J. Comparison with the previous work using this questionnaire*

A comparison between the results of Cannavò, Calandra, Pratticò, *et al.* [14] and our results is not fully possible because we only used two subtasks of their second scenario, while they asked the questionnaire after each full scenario. We removed the other subtasks for a few reasons. First, the VR shoes were not currently compatible with one of the tasks (backward walking). There were also some tasks, such as walking up the stairs and walking at heights that would not contribute to the direction results we were interested in, while including them would increase the total experiment duration. Aside from this, this task (including a large city 3D model) contained visual performance issues, causing stutters and delays in optical flow, even though the PC used had high-performing VR-ready hardware. Thereby, our questionnaire analysis was done differently, since the calculation method by Cannavò, Calandra, Pratticò, *et al.* [14] is quite involved, loses quite some statistical power in their scoring method, and calculating an overall score per condition with their method abstracts away information that could be gleaned from pairwise comparisons between conditions. Therefore, to gain more statistical power including data from all individual questions and allow for significant pairwise comparisons in both separate user experience factors and overall user experience, we used a linear mixed model instead.

### *K. Applicability of our locomotion algorithm to other locomotion techniques*

One of the goals of our locomotion algorithm was that it may be applicable to many types of full-gait locomotion techniques. Here we discuss how this may be applicable or adjusted to fit to other locomotion techniques. As discussed in Appendix A, we adjusted the locomotion speed algorithm based on a few hardware limitations of the current shoes. For the VR shoes, the speed only gradually ramps up after heel strike, see Figure 23 in Appendix A. This, combined with 150 to 300 ms latency until the standing foot started moving backward from the moment the user started moving their other foot forward, caused us to also use the lifted foot to calculate the speed. If the locomotion technique does not have these limitations, such as in a passive slidemill, or in a highly responsive and high torque active locomotion device, it might be possible to purely base it on the standing foot velocity, as described in subsection A-F in Appendix A. Using the speed in combination with the standing foot velocity direction, we believe that our algorithm would support full omnidirectional movement, even backward or sideways movement, making it suitable for any full gait technique where the foot velocity can be tracked in room space, and the user maintains the same position on average. Since we only take the horizontal component of motion, we believe it is also suitable for slidemills with a bowl as base. However, for both locomotion techniques with a surface varying in height (such as slidemill), but also recommended for any other future work, we would recommend another way than using height thresholds to detect the lifted (or standing) foot, also see subsection A-K. Thereby, if the smoothness of the tracking signal is better (no noise caused by the foot moving back, and no play in tracker motion between the foot and the tracker), the EWMA can be toned down to improve latency even further. An additional consideration is that the EWMA should also be adjusted for a differing refresh rate since it smooths per frame. Keeping this in mind, we believe this algorithm to be applicable to a wide range of techniques.

Similar to our user experience findings here, we also expect that lower body directions will be most suitable for other full-gait locomotion techniques. Similar to normal walking, we locate the goal with our heads. Once we aim for a new target, the hip, and feet align gradually as we orient towards our goal, see subsection A-H in Appendix A. If another locomotion full gait technique allows us to do the same, we expect a similar response for that technique.

### *L. Future work*

Since this work was mainly focussed on the direction, we discuss that here, and future work on the speed in subsection A-K. The future work on the VR shoes themselves is discussed in subsection IV-C.

*1) Future work on Optical Flow direction:* There was no significant UX difference between hip, average feet, and standing foot velocity directions. One explanation is that these anticipate the walking direction less in advance than the head, and therefore contrast most with the head and less with each other. Additionally, it is probably due to the lower body directions not differing a lot from each other during the experiment, which can be seen in Figure 18 in Appendix A. There, it can also be seen that as the user was looking around for their next target, the head, hip and feet directions got a larger difference in angles. In the current virtual environment, users had no reason to look off-center, aside from anticipating their path, mostly while standing still in the multi-straight-line walking. For most turns in the curved walking, only small head movements suffices. In the few steeper curves, it became more visible that the head and hip separated from the feet. However, in games, especially shooters, looking to the side will often be required to for example aim a weapon, and we expect this the users would rotate their whole upper body, and might also rotate their hips/feet with larger relative differences when doing so. Therefore, it could be that requiring participants to shoot targets to the side while walking could have a significantly different outcome on the three lower body directions that currently show no significant differences.

*2) Future work on Optical flow effect on User Experience:* We discovered some variables as significantly changing (Ease of Use, Input Responsiveness, and Appropriateness), with different optical flow algorithms, and also found some estimates (similar to the difference between averages), that were even larger, and therefore could have been influenced more than the user experience factors we found significance for. That these were not found significant is most likely due to the smaller number of questions. Therefore, more research is needed with a larger sample size to see if there is indeed a larger and significant effect on these factors.

## V. CONCLUSION

This study evaluated an algorithm mapping physical walking on VR shoes to the virtual motion displayed in the VR headset, and which direction mapping in this gave the best user experience. Thereby we were interested in which user experience factors were most impacted by differences in optical flow, and the user experience of walking on the VR shoes. We developed an algorithm to convert tracker data to optical flow, which we believe is suitable for a wide variety of full-gait locomotion techniques. The participants evaluated four direction vectors for virtual movement: head orientation, hip orientation, average feet orientation, and standing foot orientation. This resulted in the following conclusions:

- No significant differences in user experience were found between hip, average feet, and standing foot orientation, with no difference among the lower body conditions. Therefore we recommend using lower-body locomotion orientation for full-gait user-centering locomotion techniques.
- Contrary to previous works comparing optical flow directions, we found no significant difference in VR sickness, Mental Effort, and Presence. Instead, we found an effect on Ease of Use, Input Responsiveness, and Appropriateness, and indication that some other factors may be influenced even more. Further research is needed to discover the broad influence of optical flow on user experience in VR locomotion.
- While walking on the VR shoes was not experienced as completely comfortable and natural, all participants under 60 years old were able to learn to walk on the shoes within 10 minutes.

## VI. THANK YOU NOTES

This thesis has been a product of sweat and tears... and well, blood if you count the paper cuts.

First of all, dad, thank you for the magic you taught me, this is for you.

Mom, I know moms are supposed to be dedicated and loving, but I think you just broke the new record. Thank you so much for supporting me every day.

Bram, thank you for the welcome home feeling. That really saves any day. Also thank you for taking care of cooking so I did not have to eat instant noodles, and of course, the hugs!

Sandrine, I was so glad we could help each other online. It made thesis so much more fun! I will do my bestestest to support you in the final bits of your PhD. You got this!

Joseph, thank you for planting a smile on my face, especially during our time in the lab. To you, and all the other students and PhD's in Laura's lab, thank you so much for helping me out with your ideas, they helped a ton!

SenseGlove, thank you so much for providing an amazing workplace. You all returned the joy in my life, and I am happy to work with you every day. Also thank you for giving me space for my thesis, and for supplying me with some hard-to-print 3D prints when I dearly needed them!

Andrei, I still remember walking in daily with yet another broken shoe. You always went out of your way to help me repair one when I needed to just before a user test deadline. Thank you for always finding time to help!

Alexis, thank you for all the time spent with me in R and explaining me statistics. Without you, my grasp of statistics would be less than of magic, and my p-values might as well have been randomized.

Finally, Laura, you are obviously a clairvoyant superwoman. First, how is it possible that you knew exactly what I needed to hear or do at the right moment? On the superwoman part, can you please teach me your self-duplication power? I cannot phantom how you managed to live the associate professor role, support me in every way possible, give feedback on my 66-page literature thesis while running a lab with around 20 other students (also sending you their literature thesis), all while providing the best feedback on reports and in meetings I ever got. You have put the super in supervisor. This thesis would not exist, and would not be as good without you.

I might forget some very important people (my sleep was extremely limited... cough, non-existent). In case you read yourself in this statement, I'm so sorry, and thank you so much!

## REFERENCES

[1] F. Corelli, E. Battegazzorre, F. Strada, A. Bottino, and G. P. Cimellaro, "Assessing the usability of different virtual reality systems for firefighter training:" in *Proceedings of the 15th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, Valletta, Malta: SCITEPRESS - Science and Technology Publications, 2020, pp. 146–153, ISBN: 978-989-758-402-2. DOI: 10.5220/0008962401460153.

[2] L. Marchal-Crespo and R. Riener, "Technology of the robotic gait orthosis lokomat," *Neurorehabilitation Technology*, D. J. Reinkensmeyer, L. Marchal-Crespo, and V. Dietz, Eds., pp. 665–681, 2022, Springer International Publishing, ISBN: 978-3-031-08995-4. DOI: 10.1007/978-3-031-08995-4_29.

[3] T. Porssut, F. Iwane, R. Chavarriaga, *et al.*, “EEG signature of breaks in embodiment in VR,” *Public Library of Science*, vol. 18, no. 5, May 11, 2023, ISSN: 1932-6203. DOI: 10.1371/journal.pone.0282967.

[4] M. Di Luca, H. Seifi, S. Egan, and M. Gonzalez-Franco, “Locomotion vault: The extra mile in analyzing VR locomotion techniques,” in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, Yokohama Japan: ACM, May 6, 2021, pp. 1–10, ISBN: 978-1-4503-8096-6. DOI: 10.1145/3411764.3445319.

[5] E. Chang, H. T. Kim, and B. Yoo, “Virtual reality sickness: A review of causes and measurements,” *International Journal of Human–Computer Interaction*, vol. 36, no. 17, pp. 1658–1682, Oct. 20, 2020, ISSN: 1044-7318, 1532-7590. DOI: 10.1080/10447318.2020.1778351.

[6] C. N. Aldaba and Z. Moussavi, “Effects of virtual reality technology locomotive multi-sensory motion stimuli on a user simulator sickness and controller intuitiveness during a navigation task,” *Medical & Biological Engineering & Computing*, vol. 58, no. 1, pp. 143–154, Jan. 1, 2020, ISSN: 1741-0444. DOI: 10.1007/s11517-019-02070-2.

[7] A. Bashir, T. De Regt, and C. M. Jones, “Comparing a friction-based uni-directional treadmill and a slip-style omni-directional treadmill on first-time HMD-VR user task performance, cybersickness, postural sway, posture angle, ease of use, enjoyment, and effort,” *International Journal of Human-Computer Studies*, vol. 179, p. 103 101, Nov. 1, 2023, ISSN: 1071-5819. DOI: 10.1016/j.ijhcs.2023.103101.

[8] L.-O. Wehden, F. Reer, R. Janzik, W. Y. Tang, and T. Quandt, “The slippery path to total presence: How omnidirectional virtual reality treadmills influence the gaming experience,” *Media and Communication*, vol. 9, no. 1, pp. 5–16, Jan. 6, 2021, ISSN: 2183-2439. DOI: 10.17645/mac.v9i1.3170.

[9] C. N. Aldaba, P. J. White, A. Byagowi, and Z. Moussavi, “Virtual reality body motion induced navigational controllers and their effects on simulator sickness and pathfinding,” in *39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Jul. 2017, pp. 4175–4178. DOI: 10.1109/EMBC.2017.8037776.

[10] D. Calandra, F. Lamberti, and M. Migliorini, “On the usability of consumer locomotion techniques in serious games: Comparing arm swinging, treadmills and walk-in-place,” in *2019 IEEE 9th International Conference on Consumer Electronics (ICCE-Berlin)*, Sep. 2019, pp. 348–352. DOI: 10.1109/ICCE-Berlin47944.2019.8966165.

[11] K. Jochymczyk-Woźniak, K. Nowakowska, J. Polechoński, S. Sładczyk, and R. Michnik, “Physiological gait versus gait in VR on multidirectional treadmill-comparative analysis,” *Medicina (Lithuania)*, vol. 55, no. 9, 2019, ISSN: 1010-660X. DOI: 10.3390/medicina55090517.

[12] S. Pyo, H. Kim, and J. Yoon, “Control scheme for sideways walking on a user-driven treadmill,” in *2022 International Conference on Robotics and Automation (ICRA)*, May 2022, pp. 6963–6968. DOI: 10.1109/ICRA46639.2022.9812403.

[13] R. P. Darken, William R. Cockayne, W. Cockayne, and D. Carmein, “The omni-directional treadmill: A locomotion device for virtual worlds,” pp. 213–221, Oct. 1, 1997. DOI: 10.1145/263407.263550.

[14] A. Cannavò, D. Calandra, F. G. Pratticò, V. Gatteschi, and F. Lamberti, “An evaluation testbed for locomotion in virtual reality,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 27, no. 3, pp. 1871–1889, Mar. 2021, ISSN: 1941-0506. DOI: 10.1109/TVCG.2020.3032440.

[15] C. T. Tan, L. C. Foo, A. Yeo, *et al.*, “Understanding user experiences across VR walking-in-place locomotion methods,” in *CHI Conference on Human Factors in Computing Systems*, New Orleans LA USA: ACM, Apr. 29, 2022, pp. 1–13, ISBN: 978-1-4503-9157-3. DOI: 10.1145/3491102.3501975.

[16] P. Schmitz, J. Hildebrandt, A. C. Valdez, L. Kobbelt, and M. Ziefle, “You spin my head right round: Threshold of limited immersion for rotation gains in redirected walking,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 24, no. 4, pp. 1623–1632, Apr. 2018, ISSN: 1077-2626. DOI: 10.1109/TVCG.2018.2793671.

[17] F. Steinicke, Y. Visell, J. Campos, and A. Lécuyer, *Human walking in virtual environments: Perception, technology, and applications* (Human Walking in Virtual Environments: Perception, Technology, and Applications). New York, NY: Springer, 2013, vol. 9781441984326, 402 pp., ISBN: 978-1-4419-8432-6. DOI: 10.1007/978-1-4419-8432-6.

[18] J. Reason and J. J. Brand, *Motion sickness*. the University of Michigan: Academic Press, 1975, 310 pp., ISBN: 978-0-12-584050-7.

[19] C. Zanbaka, S. Babu, D. Xiao, A. Ulinski, L. F. Hodges, and B. Lok, “Effects of travel technique on cognition in virtual environments,” in *IEEE Virtual Reality 2004*, IEEE, 2004, pp. 149–286. DOI: 10.1109/VR.2004.1310068.

[20] C.-G. Lee and O. Kwon, “Identification of the difference threshold for curvature gain of redirected walking,” *Virtual Reality*, Feb. 6, 2023, ISSN: 1359-4338, 1434-9957. DOI: 10.1007/s10055-023-00763-6.

[21] G. Bruder, P. Lubos, and F. Steinicke, “Cognitive resource demands of redirected walking,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 21, no. 4, pp. 539–544, Apr. 18, 2015, ISSN: 1077-2626, 1941-0506, 2160-9306. DOI: 10.1109/TVCG.2015.2391864.

[22] N. C. Nilsson, S. Serafin, and R. Nordahl, “Walking in place through virtual worlds,” *18th International Conference on Human-Computer Interaction*, Lecture Notes in Computer Science, pp. 37–48, 2016, Publisher: Springer. DOI: 10.1007/978-3-319-39516-6_4.

[23] Y.-J. Li, F. Steinicke, and M. Wang, "A comprehensive review of redirected walking techniques: Taxonomy, methods, and future directions," *Journal of Computer Science and Technology*, vol. 37, no. 3, pp. 561–583, Jun. 2022, ISSN: 1000-9000, 1860-4749. DOI: 10.1007/s11390-022-2266-7.

[24] C. Cortes, H.-T. Chen, and C.-T. Lin, "Analysis of VR sickness and gait parameters during non-isometric virtualwalking with large translational gain," presented at the Proceedings - VRCAI 2019: 17th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and its Applications in Industry, 2019, ISBN: 978-1-4503-7002-8. DOI: 10.1145/3359997.3365694.

[25] "VIRTUSPHERE — discover virtual reality." http://www.virtusphere.com/. (), [Online]. Available: http://www.virtusphere.com/ (visited on 03/31/2022).

[26] E. Medina, R. Fruland, and S. Weghorst, "VIRTUSPHERE: Walking in a human size VR "hamster ball"," presented at the Proceedings of the Human Factors and Ergonomics Society, vol. 3, 2008, pp. 2102–2106. DOI: 10.1177/154193120805202704.

[27] M. Nabiyouni, A. Saktheeswaran, D. A. Bowman, and A. Karanth, "Comparing the performance of natural, semi-natural, and non-natural locomotion techniques in virtual reality," in *2015 IEEE Symposium on 3D User Interfaces (3DUI)*, Mar. 2015, pp. 3–10. DOI: 10.1109/3DUI.2015.7131717.

[28] H. Iwata, H. Yano, and F. Nakaizumi, "Gait master: A versatile locomotion interface for uneven virtual terrain," presented at the Proceedings - Virtual Reality Annual International Symposium, 2001, pp. 131–137. DOI: 10.1109/VR.2001.913779.

[29] M. Reese (Holotron). "Holotron." Bielefeld, Germany. (), [Online]. Available: https://www.holotron.com/ (visited on 06/03/2022).

[30] H. Iwata, H. Yano, H. Fukushima, and H. Noma, "CirculaFloor [locomotion interface]," *IEEE Computer Graphics and Applications*, vol. 25, no. 1, pp. 64–67, Jan. 2005, ISSN: 1558-1756. DOI: 10.1109/MCG.2005.5.

[31] VR Lab University of Tsukuba. "Robot tile 2." (Aug. 11, 2016), [Online]. Available: https://www.youtube.com/watch?v=oO4VTUbzEL0 (visited on 07/20/2023).

[32] S. Pyo, H. Lee, and J. Yoon, "Development of a novel omnidirectional treadmill-based locomotion interface device with running capability," *Applied Sciences*, vol. 11, p. 4223, May 6, 2021. DOI: 10.3390/app11094223.

[33] B. Freeman. "Infinadeck," Infinadeck. (2015), [Online]. Available: https://www.infinadeck.com/ (visited on 11/12/2021).

[34] ACTVR. "ACTVR - 360 movement in virtual reality." (2020), [Online]. Available: https://www.actvr.com/ (visited on 04/14/2022).

[35] "Omnideck: Redefining VR treadmills," Omnifinity. (), [Online]. Available: https://www.omnifinity.se/ (visited on 04/20/2022).

[36] MIXED. "StriderVR virtual-reality-treadmill walking test." (Aug. 13, 2017), [Online]. Available: https://www.youtube.com/watch?v=_b3VwisaYic (visited on 06/26/2023).

[37] R. P. McMahan, C. Lai, and S. K. Pal, "Interaction fidelity: The uncanny valley of virtual reality interactions," Book Title: Virtual, Augmented and Mixed Reality, vol. 9740, Cham: Springer International Publishing, 2016, pp. 59–70, ISBN: 978-3-319-39906-5. DOI: 10.1007/978-3-319-39907-2_6.

[38] R. Ruddle, E. Volkova, and H. BüLthoff, "Walking improves your cognitive map in environments that are large-scale and large in extent," *ACM Transactions on Computer-Human Interaction*, vol. 18, no. 2, 2011, ISSN: 1557-7325. DOI: 10.1145/1970378.1970384.

[39] ACTVR. "Omnidirectional treadmill CM1 - running in virtual reality - work in progress 2." (Nov. 29, 2021), [Online]. Available: https://www.youtube.com/watch?v=R-bWKF6ySTo (visited on 04/14/2022).

[40] Finally Functional. "Cheap DIY VR slidemill setup." (Apr. 16, 2023), [Online]. Available: https://www.youtube.com/watch?v=tuN-QA6cY7Y (visited on 06/26/2023).

[41] Cybershoes. "Cybershoes for quest & SteamVR VR gaming shoe." https://www.cybershoes.com. (2020), [Online]. Available: https://www.cybershoes.com (visited on 11/12/2021).

[42] B. Factor. "EKTO VR: Virtual reality training with robotics," Ekto VR. https://ektovr.com/. (2018), [Online]. Available: https://ektovr.com/ (visited on 11/26/2021).

[43] A. Foxcroft. "Freeaim VR shoes," Freeaim VR shoes. United Kingdom, https://www.freeaim.com. (2021), [Online]. Available: https://www.freeaim.com (visited on 11/12/2021).

[44] H. Iwata, H. Yano, and H. Tomioka, "Powered shoes," in *ACM SIGGRAPH 2006 Emerging technologies on - SIGGRAPH '06*, Boston, Massachusetts: ACM Press, 2006, p. 28, ISBN: 978-1-59593-364-5. DOI: 10.1145/1179133.1179162.

[45] N. Rathi and M. Ratner, "Augmented and/or virtual reality footwear," pat. WO2018208924A1, Nov. 15, 2018.

[46] Finally Functional. "Virtual reality motorized shoe build log #32." (Jan. 28, 2021), [Online]. Available: https://www.youtube.com/watch?v=7pQpHFiUKTw (visited on 06/26/2023).

[47] Skarredghost. "Ekto VR hands-on: The craziest boots to walk in VR," The Ghost Howls. (Nov. 25, 2021), [Online]. Available: https://skarredghost.com/2021/11/25/ekto-vr-hands-on-review/ (visited on 07/11/2023).

[48] M. Nabiyouni and D. A. Bowman, "A taxonomy for designing walking-based locomotion techniques for virtual reality," in *Proceedings of the 2016 ACM Companion on Interactive Surfaces and Spaces*, ser. ISS '16 Companion, New York, NY, USA: Association for Computing Machinery, Nov. 6, 2016, pp. 115–121, ISBN: 978-1-4503-4530-9. DOI: 10.1145/3009939.3010076.

[49] T. Banton, J. Stefanucci, F. Durgin, A. Fass, and D. Proffitt, “The perception of walking speed in a virtual environment,” *Presence: Teleoperators and Virtual Environments*, vol. 14, no. 4, pp. 394–406, Aug. 1, 2005. DOI: 10.1162/105474605774785262.

[50] I. Choi, J.-J. Park, S. Kim, and H.-C. O. Li, “Effect of inconsistency between visually perceived walking speed and physically perceived walking speed on VR sickness in VR-treadmill walking,” *Science of Emotion and Sensibility*, vol. 23, no. 3, pp. 79–90, 2020, ISSN: 1226-8593. DOI: 10.14695/KJSOS.2020.23.3.79.

[51] N. C. Nilsson, S. Serafin, and R. Nordahl, “Establishing the range of perceptually natural visual walking speeds for virtual walking-in-place locomotion,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 20, no. 4, pp. 569–578, Apr. 1, 2014. DOI: 10.1109/tvcg.2014.21.

[52] T. Perrin, H. A. Kerhervé, C. Faure, A. Sorel, B. Bideau, and R. Kulpa, “Enactive approach to assess perceived speed error during walking and running in virtual reality,” in *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, Mar. 2019, pp. 622–629. DOI: 10.1109/VR.2019.8798209.

[53] F. Steinicke, G. Bruder, J. Jerald, H. Frenz, and M. Lappe, “Estimation of detection thresholds for redirected walking techniques,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 16, no. 1, pp. 17–27, Jan. 2010, ISSN: 1941-0506. DOI: 10.1109/TVCG.2009.62.

[54] R. Reed-Jones, J. Reed-Jones, L. A. Vallis, and M. Hollands, “The effects of constraining eye movements on visually evoked steering responses during walking in a virtual environment,” *Experimental Brain Research*, vol. 197, no. 4, pp. 357–367, Aug. 2009, ISSN: 1432-1106. DOI: 10.1007/s00221-009-1923-1.

[55] M. Al Zayer, P. MacNeilage, and E. Folmer, “Virtual locomotion: A survey,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 26, no. 6, pp. 2315–2334, 2020, ISSN: 1077-2626. DOI: 10.1109/TVCG.2018.2887379.

[56] N. Nilsson, S. Serafin, M. Laursen, K. Pedersen, E. Sikstrom, and R. Nordahl, “Tapping-in-place: Increasing the naturalness of immersive walking-in-place locomotion through novel gestural input,” presented at the IEEE Symposium on 3D User Interface 2013, 3DUI 2013 - Proceedings, 2013, pp. 31–38, ISBN: 978-1-4673-6097-5. DOI: 10.1109/3DUI.2013.6550193.

[57] J. N. Templeman, L. E. Sibert, R. C. Page, and P. S. Denbrook, “Immersive simulation to train urban infantry combat,” Technical report, DTIC Document, 2006, Section: Technical Reports.

[58] R. A. Ruddle and S. Lessels, “The benefits of using a walking interface to navigate virtual environments,” *ACM Transactions on Computer-Human Interaction*, vol. 16, no. 1, 5:1–5:18, Apr. 23, 2009, ISSN: 1073-0516. DOI: 10.1145/1502800.1502805.

[59] B. Williams, M. McCaleb, C. Strachan, and Y. Zheng, “Torso versus gaze direction to navigate a VE by walking in place,” presented at the Proceedings - SAP 2013: ACM Symposium on Applied Perception, 2013, pp. 67–70, ISBN: 978-1-4503-2262-1. DOI: 10.1145/2492494.2492512.

[60] L. Joseph and K. Ernst, “8.5.1 spatial steering techniques,” in *3D User Interfaces: Theory and Practice*, Apr. 7, 2024, ISBN: 978-0-13-403432-4.

[61] L. Rebenitsch and D. Engle, “The effects of steering locomotion on user preference and accuracy in virtual environments,” *Presence*, vol. 28, pp. 153–168, Jan. 2019, Conference Name: Presence, ISSN: 1054-7460. DOI: 10.1162/pres_a_00345.

[62] A. Elferink, *AmberElferink/VRshoes-UX-unity-environment*, Jun. 8, 2024.

[63] “VIVE tracker (3.0),” HTC Vive Tracker 3.0. HTC, Taiwan. (2021), [Online]. Available: https://www.vive.com/eu/accessory/tracker3/ (visited on 04/14/2024).

[64] “Vive pro series.” HTC, Taiwan. (2018), [Online]. Available: https://www.vive.com/eu/product/ (visited on 06/08/2024).

[65] R. S. Kalawsky, “VRUSE—a computerised diagnostic tool: For usability evaluation of virtual/synthetic environment systems,” *Applied Ergonomics*, vol. 30, no. 1, pp. 11–25, Feb. 1999, ISSN: 00036870. DOI: 10.1016/S0003-6870(98)00047-7.

[66] I. O. for Standardization, *ISO 9241-400: Ergonomics of human — system interaction — part 400: Principles and requirements for physical input devices*, 2007.

[67] R. S. Kennedy, N. E. Lane, K. S. Berbaum, and M. G. Lilienthal, “Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness,” *The International Journal of Aviation Psychology*, vol. 3, no. 3, pp. 203–220, Jul. 1, 1993, ISSN: 1050-8414. DOI: 10.1207/s15327108ijap0303_3.

[68] D. Lochner and J. Gain, “VR natural walking in impossible spaces,” presented at the Proceedings - MIG 2021: 14th ACM SIGGRAPH Conference on Motion, Interaction, and Games, ZSCC: 0000000, 2021, ISBN: 978-1-4503-9131-3. DOI: 10.1145/3487983.3488305.

[69] J. Häkkinen, F. Ohta, T. Kawai, F. Ohta, and T. Kawai, “Time course of sickness symptoms with HMD viewing of 360-degree videos,” *Electronic Imaging*, vol. 31, pp. 1–11, Nov. 1, 2018, Publisher: Society for Imaging Science and Technology, ISSN: 2470-1173. DOI: 10.2352/J.ImagingSci.Technol.2018.62.6.060403.

[70] Y.-H. Peng, C. Yu, S.-H. Liu, *et al.*, “WalkingVibe: Reducing virtual reality sickness and improving realism while walking in VR using unobtrusive head-mounted vibrotactile feedback,” in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, ser. CHI ’20, New York, NY, USA: Association for Computing Machinery, Apr. 21, 2020, pp. 1–12, ISBN: 978-1-4503-6708-0. DOI: 10.1145/3313831.3376847.

[71] K. Uliano, E. Lambert, R. Kennedy, and D. Sheppard, “The effects of asynchronous visual delays on simulator flight performance and the development of simulator sickness symptomatology,”

Technical report, DTIC Document, 1986, URL: https://apps.dtic.mil/sti/citations/ADA180196.

[72] M. E. McCauley and T. J. Sharkey, "Cybersickness: Perception of self-motion in virtual environments," *Presence: Teleoperators and Virtual Environments*, vol. 1, no. 3, pp. 311–318, Aug. 1, 1992. DOI: 10.1162/pres.1992.1.3.311.

[73] J. C. Selinger, J. L. Hicks, R. W. Jackson, C. M. Wall-Scheffler, D. Chang, and S. L. Delp, "Running in the wild: Energetics explain ecological running speeds," *Current Biology*, vol. 32, no. 10, 2309–2315.e3, May 23, 2022, ISSN: 0960-9822. DOI: 10.1016/j.cub.2022.03.076.

[74] J. Feasel, M. Whitton, and J. Wendt, "LLCM-WIP: Low-latency, continuous-motion walking-in-place," presented at the 3DUI - IEEE Symposium on 3D User Interfaces 2008, 2008, pp. 97–104, ISBN: 978-1-4244-2047-6. DOI: 10.1109/3DUI.2008.4476598.

[75] M. Usoh, K. Arthur, M. C. Whitton, *et al.*, "Walking > walking-in-place > flying, in virtual environments," in *Proceedings of the 26th annual conference on Computer graphics and interactive techniques - SIGGRAPH '99*, Not Known: ACM Press, 1999, pp. 359–364, ISBN: 978-0-201-48560-8. DOI: 10.1145/311535.311589.

[76] M. Slater, M. Usoh, and A. Steed, "Taking steps: The influence of a walking technique on presence in virtual reality," *ACM Transactions on Computer-Human Interaction*, vol. 2, no. 3, pp. 201–219, Sep. 1, 1995, ISSN: 1073-0516. DOI: 10.1145/210079.210084.

[77] J. Lee, S. C. Ahn, and J.-I. Hwang, "A walking-in-place method for virtual reality using position and orientation tracking," *Multidisciplinary Digital Publishing Institute: Sensors*, vol. 18, no. 9, p. 2832, Sep. 2018. DOI: 10.3390/s18092832.

[78] S. Tregillus and E. Folmer, "VR-STEP: Walking-in-place using inertial sensing for hands free navigation in mobile VR environments," in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, San Jose California USA: ACM, May 7, 2016, pp. 1250–1255, ISBN: 978-1-4503-3362-7. DOI: 10.1145/2858036.2858084.

[79] Deca. "DecaMove," Deca. (2021), [Online]. Available: https://www.deca.net/decamove/ (visited on 04/29/2022).

[80] J. N. Templeman, P. S. Denbrook, and L. E. Sibert, "Virtual locomotion: Walking in place through virtual environments," *Presence*, vol. 8, no. 6, pp. 598–617, Dec. 1999, ISSN: 1054-7460. DOI: 10.1162/105474699566512.

[81] W. T. Welford, J. M. T. Brebner, and N. Kirby, *Reaction Times*. Stanford University, 1980, 132 pp.

[82] N. Sekiya and H. Nagasaki, "Reproducibility of the walking patterns of normal young adults: Test-retest reliability of the walk ratio(step-length/step-rate)," *Gait & Posture*, vol. 7, no. 3, pp. 225–227, May 1, 1998, ISSN: 1879-2219. DOI: 10.1016/s0966-6362(98)00009-5.

# Appendix A
## Locomotion algorithm design decisions

In this appendix, we will discuss all design steps, tests, and findings for the locomotion speed used. First, we some previous works on the transfer function, specifically on locomotion speed, are discussed in subsection A-A. We show the steps of designing our speed algorithm, subsection A-B. This initial design was then tuned and tested on the resulting locomotion speed subsection A-G. After these tests, we implemented the algorithm into Unity for which we mention changes here subsection A-I. Then, we discuss what we think caused the results indicating a slower experienced locomotion speed from the user test in subsection A-J, and provide recommendations for future improvements on the speed algorithm in subsection A-K. We also briefly discuss a plot for directions over time in subsection A-H.

### *A. Previous works on transfer function speed*

Transfer functions have as far as we have found, not been explicitly described for full gait locomotion techniques, however, multiple options have been described for walk-in-place (WIP). There are many different designs, but most use the following stages: step detection or action recognition, virtual speed calculation, virtual direction calculation and ignoring actions and maneuvering.

An overview of all WIP transfer function algorithms is given in Table II, and the related speed profiles in Figure 10. Most works use the VR headsets (HMD) position, velocity or acceleration for their algorithm. Papers using speed, used numeric differentiation of the positions to obtain the speeds [74]–[76]. In Table II, some values are noted with a star, which means their data is possibly not correct or incomparable, see subsection A-L.

TABLE II

Walk-in-place methods comparison. This table shows the input data used to calculate the speed ($x$ = position x, $dx$ = velocity x, $ddx$ = acceleration of $x$ with respect to the physical room, y-axis is up, z forward and x sideways), the accuracy of a step detection (FP = % false positives and % FN = false negative detections), latency (time between start and stop of walking and the resulting VR motion), directional: the tracker rotation used to steer VR motion, and the speed profile used, see Figure 10. $*$ = See discussion: possibility of incorrect value or discussion point.

| **Paper** | **Input data** | **Accuracy (mean)** | **Latency** | **Directional** | **Speed profile** |
|---|---|---|---|---|---|
| Lee et al., 2018 | HMD<br>($Y_{pos}, X_{rot}, Y_{rot}$) | 99.32%* | Start: 1-2* steps + 44ms<br>Stop: 44ms | HMD $Y_{rot}$ | Saw-tooth |
| Usoh et al., 1999 | HMD<br>($dx, dy, dz$) | >91% time*<br>FP: 3%, FN: 32% | 500 ms | HMD $Y_{rot}$ | Impulse |
| Slater et al., 1995 | HMD<br>($dx, dy, dz$) | 91% of time<br>FP: 10%, FN: 16% | - | HMD $Y_{rot}$ | Impulse |
| Tregillus & Folmer, 2016 | HMD<br>($ddy$) | - | 100-200ms | HMD $Y_{rot}$ | Saw-tooth |
| Feasel et al, 2008 | Heels<br>($dy$) | - | 88-156ms | Chest $Y_{rot}$ | Direct tracker speed to locomotion speed |

| **Paper** | **Ignores false positives:** |
|---|---|
| Lee et al., 2018 | Step-in-place (only jogging accepted), maneuvering, crouching, looking-up-down |
| Usoh et al., 1999 | Not mentioned, but trained neural network to ignore bending down, moving around, turning the head, and mixtures of these. |
| Slater et al., 1995 | - |
| Tregillus & Folmer, 2016 | - |
| Feasel et al, 200 | Subtracting small heel speed and clamping the result to allow small heel speeds in maneuvering |

Some methods detect each step, and afterward, confirm by looking at the vertical displacement of the feet and time since the last step whether the step was valid [77], [78]. Some do not detect individual steps, but rather if the action in this frame is WIP or not [75], [76] using a neural network, adding to the speed by impulse if this is the case (while the speed is reduced by artificial friction each frame to slow speed). Feasel, Whitton, and Wendt [74] does not register discrete steps but maps the speed of the heels to locomotion speed. They focus on creating continuous locomotion output, for which the speed is controllable even within a step. Thereby they achieved low latency, 115 to 165 ms latency (which is a fraction of a step) on start, and 88 to 110 ms latency on stop depending on walking speed [74] versus 500 ms [75], or 1 to 2 steps + 44ms on start and 44 ms on stop [77]. For our algorithm, we use the strategy of Feasel, Whitton, and Wendt [74], since they showed low latency for both starting and stopping, reliable output (no extra option for false positive/negative step detection), and a responsive and direct relation between the user's actions and speed. Thereby, we expect this to work well for full-gait locomotion since there might be a way to map the speed 1:1 with the real speed of the shoes to virtual motion.

For the direction condition they used head orientation [75]–[78], tracker on the waist [15], [79], chest [74], hands, or using the joystick for turning [80]. We describe further research on directions in subsubsection I-B2.

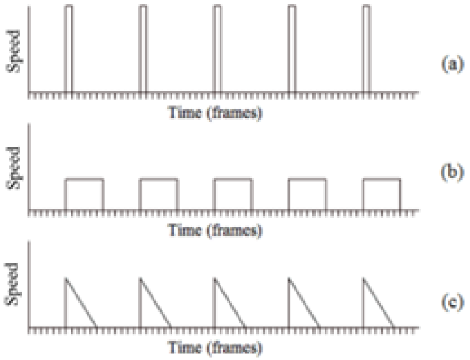


Fig. 10. Virtual speed profiles of previous works. a) Impulse: the full speed on the frame of detected input only; B) block: The velocity is spread out over multiple frames, causing an abrupt stop after step; c) Saw-tooth movement: Full speed in the first frame, with the decelerated speed in subsequent frames. The start of each pulse is the step detection. Adapted from [74].

Some of these previous works also ignore some actions to prevent false step detection, such as substracting a small constant velocity from the filtered heel velocity, and then clamping the values. This allowed for small heel movements to be ignored [74]. Nilsson, Serafin, Laursen, *et al.* [56] ignore normal room-scale walking by only recognizing WIP if the horizontal step length was under 0.3m. Since the VR shoes themselves do not currently support maneuvering, we did not pursue this in our design.

### *B. Designing the speed algorithm*

In this section, we describe the process of creating the speed algorithm. First we go over the requirements subsection A-C, then the design process used to prototype our algorithm. In subsection A-F we discuss the steps taken to come to our result, first discussing converting to speed and filtering, followed by a description of what we had in mind, and why we diverted from this plan, resulting in the inclusion of the lifted foot to the speed calculation. All raw Google Sheets files including plots discussed, can be found on Github [62].

### *C. Designing the speed algorithm: Requirements*

For our speed algorithm, we had multiple goals:

- Low latency for starting and stopping, being at least under the human perception reaction time of light (190 ms) [81].
- Close to a 1:1 walking speed or at least within the detection threshold (50% of people could notice) of a factor between 0.86 and 1.26 of the user's actual walking speed [53].
- A smooth virtual locomotion speed in optical flow over time, similar to the optical flow for normal walking.
- Using the speed/velocity so it could either be used with positional trackers by derivation, or IMU's by integration of their acceleration.

We tested each of these in subsection A-G.

### *D. Designing the speed algorithm: Design process*

We performed a brief pilot with a single subject, walking both with VR shoes turned on (to convert this tracker data to locomotion speed), and with VR shoes turned off to obtain natural locomotion speed through the room as our goal, see Figure 16. To iterate algorithm designs quickly, we loaded this raw tracking data into Google Sheets (similar to Microsoft Excel). Here, multiple algorithms were iteratively implemented and plotted to see the results, using the strategy of Feasel, Whitton, and Wendt [74] as a starting point: calculating speeds, filtering the speeds, and combining them to obtain the final locomotion speed. For now, we ignore the lifted foot velocity calculation, since the lifted foot velocity was introduced later in the design stage.

### *E. Designing the speed algorithm: Velocity to speed and filtering*

While Feasel, Whitton, and Wendt [74] used the vertical heel speed, in full gait we have the horizontal component available to us, which should in theory be convertible to reach 1:1 speed. To use this in a similar algorithm, we therefore wanted to convert the velocity to a scalar representing horizontal motion, ($s_{l/r}$). To be able to plot our results, we kept this scalar signed, so we could distinguish between the foot moving forward in the lifted phase and backward when the shoe was driving the user backward in our plots. The process to obtain $s_{l/r}$ is described in Figure 3 and Equation 2.

First, just as Feasel, Whitton, and Wendt [74] did, we found significant noise, see Figure 11. We believe this noise partially comes from noise in the differentiation itself: a small time or positional tracking noise becoming more apparent on differentiation for a small timestep. Second, we think the tracker could wiggle slightly on the mount as well as the shoe relative to the foot while moving, causing additional 'noise' in the velocity. It seems this is also largely caused by the VR shoe vibrations when they are active, since while walking with the VR shoes off, this reduces the noise significantly, see subsubsection A-G3. To smooth this noise out, we employed an exponentially weighted moving average EWMA filter (exponentially weighted average filter), because it was fast to implement and seemed to work well. Just like the filter in Feasel, Whitton, and Wendt [74], using a smoothing filter introduces latency, the more smoothed the more latency, as we discuss in subsubsection A-G1.

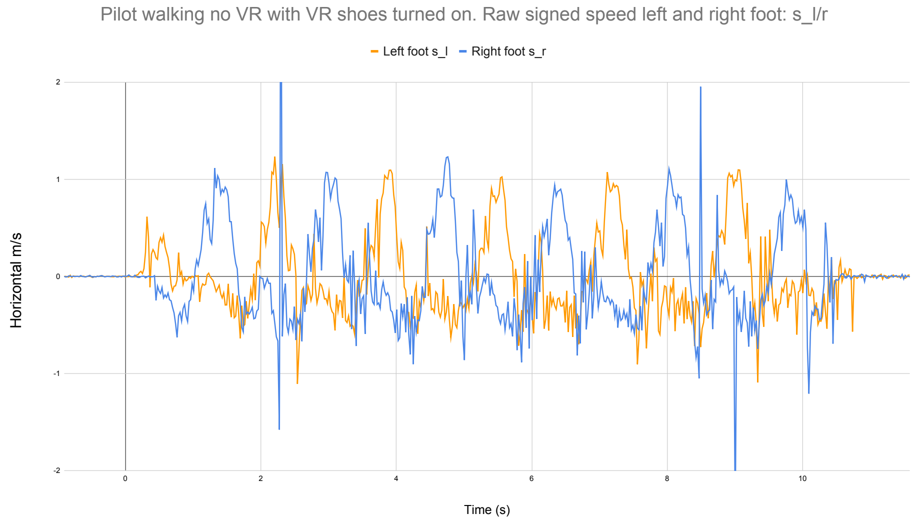


Fig. 11. This is the raw horizontal speed ($s_{l/r}$) obtained from the derivation of tracker data while walking on the VR shoes. This was recorded during an early pilot with the no VR headset covering the eyes. The speed is along the horizontal axis the participant was facing.

### *F. Designing the speed algorithm: Locomotion speed calculation as we had in mind, and the decision to include lifted foot*

*1) Our initial plan:* At first, we were planning to use EWMA speed of only the standing foot to calculate our locomotion speed, since in normal walking, the standing foot pushes off against the floor, and if we convert any velocity from room-scale frame (stationary user) to the virtual reference frame (the user is not stationary), one should subtract this standing foot velocity from all velocities so the standing foot is stationary with the virtual ground, and the rest of the body moves forwards with this relative velocity. We thought that is this speed would be combined with the standing foot velocity direction as the direction component of the velocity vector, that this would result in a natural optical flow.

*2) Execution of the plan, and findings:* Therefore, we calculated the respective EWMA for the absolute value of the standing foot $s_{standing}$, making it a normal positive speed instead of a possibly signed speed. This is described in Equation 3 for left/right foot, which it is a similar process for the standing foot. However, when we calculated this $EWMA_{standing}$, we obtained the following plot Figure 12 where we averaged the EWMA of shoes driving backward.

This showed that the resulting locomotion speed was not smooth for our standing foot attempts, probably because the VR shoes gradually ramped up in speed after the heel strike. We continued to plot some options and found that for the $EWMA_{lifted}$ the phase counters that of the standing foot, see Figure 13, which means that combining these would smooth out the locomotion speed. Additionally, the VR shoe of the standing foot would only start moving backward after 150 to 300 ms (looking at $s_{l/r}$), which would add as additional latency. Incorporating the lifted foot would result in a near immediate response when the user would start to swing their foot forward, ignoring this latency of the VR shoe on the standing foot. Therefore, we decided to combine the speed of both feet in our algorithm.

*3) Lifted foot velocity correction:* A simple average between the speeds of both feet would seem suitable, but since the standing foot drives the entire user back, including the lifted foot, this would result in a lifted foot velocity of near zero. Therefore, we opted to cancel the standing foot velocity for the lifted foot by subtracting it (resulting in the lifted foot velocity in the non-stationary frame), with $\vec{v}_{lifted} = \vec{v}_{lifted\_tracker}$ - $\vec{v}_{standing}$, also see step 1 in Figure 3. In theory, over time, $\vec{v}_{lifted}$ + $\vec{v}_{standing}$ cancel out if the user stays in the same spot in the room. However, within each step, these velocities differ from each other since they are in a counter phase, as visible in Figure 13.

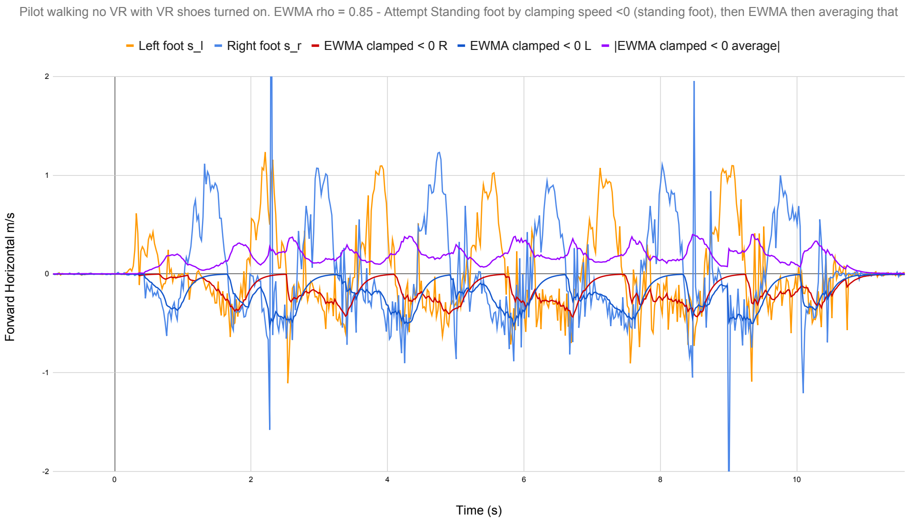


Fig. 12. This was a second attempt to calculate locomotion speed based on the standing foot. It applies the averaging filter on the individual foot speed below (standing foot) zero and then averages this.

*4) Combination of both feet:* Therefore both $EWMA_{l/r}$ were combined (should yield the same result as combined $EWMA_{lifted/standing}$)to obtain the locomotion speed, see Equation 4, which also seemed to be much smoother and with similar latency compared to the locomotion speed curve by Feasel, Whitton, and Wendt [74], and comparable to natural walking see subsubsection A-G3.

### G. Tuning and testing the speed algorithm

We tested our algorithm for each requirement in subsection A-C. First we tuned the smoothing filter to limit latency while still smoothing the data sufficiently, see subsubsection A-G1. Then, we compared the average speeds of the final locomotion speeds and raw locomotion speeds, to see if this was within the detection thresholds, see subsubsection A-G2. To see how the locomotion speed curve compared to normal walking, we compared it in subsubsection A-G3. The resulting latency was analyzed in subsubsection A-G4.

*1) EWMA $\rho$ tuning:* The EWMA smoothing can be tuned with the $\rho$. By setting it higher, the signal becomes more smooth, but this comes at the cost of latency, including starting/stopping latency. Therefore we tried out several values, see Figure 15, finally deciding on 0.85 for the final EWMA, giving a good balance between the smoothness of the signal and starting/stopping latency. An additional note is that this $\rho$ might have to be tuned depending on the frame rate and input tracking smoothness for another application.

*2) Comparison in speeds:* There were multiple artificial steps in this process, mainly the smoothing, and averaging together with the lifted foot. To check if the speed was still matching the actual walking speed, we compared multiple ways of averaging the raw unfiltered $s_l/r$ and locomotion speed/EWMA's, see Table III. This showed that the final speed was only slightly higher than the raw average speed with a factor 1.08, which is well within the detection threshold of 1.26 [53] and therefore usable for our purposes. The increase in speed mainly happens within the step of averaging the lifted and standing foot together. Thereby there is barely an influence from tuning the EWMA on the average speed, even on the extremes of 0.1 and 0.95. Therefore EWMA only changes the smoothing of the motion, and not the average speed.

*3) Comparison with natural speed plot:* To compare the smoothness of our locomotion speed with normal walking, within the brief pilot, we also recorded data walking at room-scale with the shoes off, with the VR headset on the forehead to measure the head position while walking on the VR shoes in a straight line through the room, see Figure 16. Since this also includes head motion, we added the head speed to our final locomotion speed, see Figure 17. Although the locomotion speed output is slightly more variable, it seems to follow a similar speed curve compared to normal walking, with only a slight shift in the maximum speed in the gait cycle being just before the double stance phase instead of in the double stand phase, and an otherwise relatively smooth movement, while still being responsive in starting and stopping. From Figure 17, it is also visible that before stopping on the VR shoes, the user has to shift their weight backward to compensate for their inertia continuing to move after the VR shoes stop, which is not the case for normal walking.

*4) Speed Responsiveness:* As a brief test of optical flow speed responsiveness, we took 10 random samples of the user test data, spread over different conditions/users. We measured the time from when the raw signal $\vec{s}_{l\_r}$ first becomes positive, indicating the foot starts being moved forwards/lifted, until the locomotion speed was at least 1/4 of the average locomotion

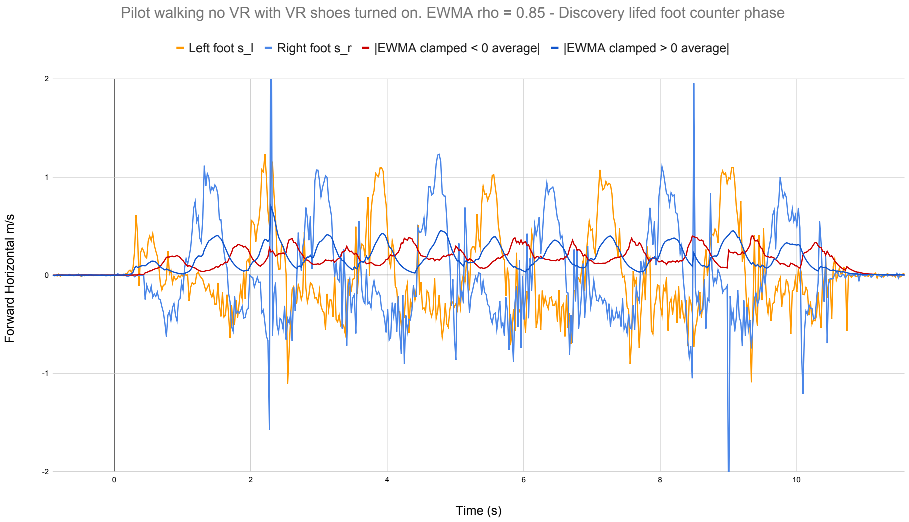


Fig. 13. Plotting the result in Figure 12 also for the lifted foot, showed that these values had a shifted phase, meaning that averaging these would result in a smoother final locomotion speed. Additionally, while the VR shoe (standing foot) takes time before it responds and drives backward, the lifted foot signal reacts nearly immediately when the user starts to swing their foot forward in a step.

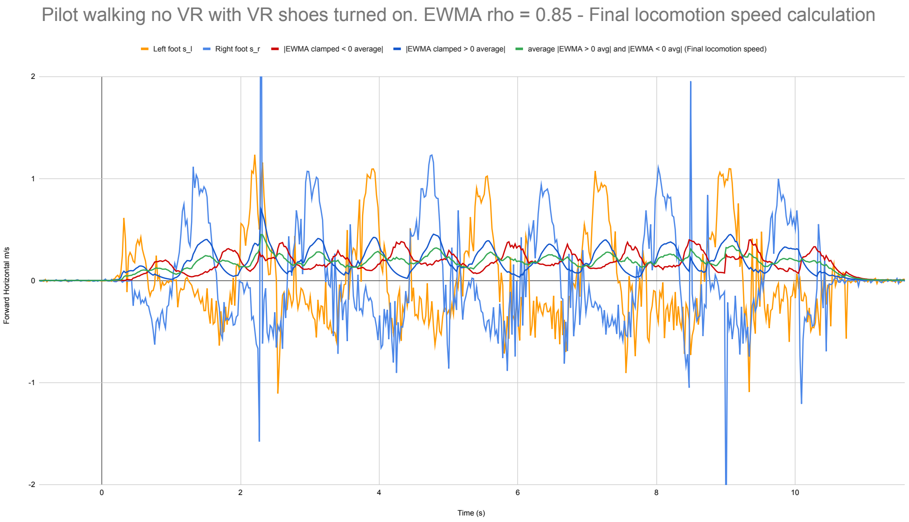


Fig. 14. This figure shows the calculation of the final locomotion speed, based on averaging the results displayed in Figure 13. While the idea came by looking at the standing foot and lifted foot separately, this is mathematically exactly the same as averaging the right and left foot, which is done in our final algorithm in Equation 4.

speed. These latencies were all under 130 ms. This is more responsive than could be obtained from basing speed purely on the standing foot velocity since the VR shoe has a latency of 150 to 300 ms until the standing shoe starts driving (measured similarly based on the raw signed tracker speeds). Thereby it is well within the human perception reaction time of light (190 ms) [81], and similar to the latency of Feasel, Whitton, and Wendt [74] (88 to 156 ms).

### *H. Designing the direction algorithm: Direction result*

How the directions were chosen was explained in subsection I-D. How to calculate the directions was explained in subsubsection II-B1. In Figure 18, we show an example of how these direction conditions differed from each other in the user test. Since the active condition was standing foot, the head can turn independently. When the user was standing still, the standing foot switched between feet rapidly (which was not noticeable to the user, since their speed was 0). Therefore we can see when the user was standing still by these horizontal broad green bands.

The first part is the multi-straight line walking task. The user starts at angle 0, which flips back and forth to 360 degrees. After the user turns around, we can see the data more clearly. First, the head searches and orients towards a new target. Once

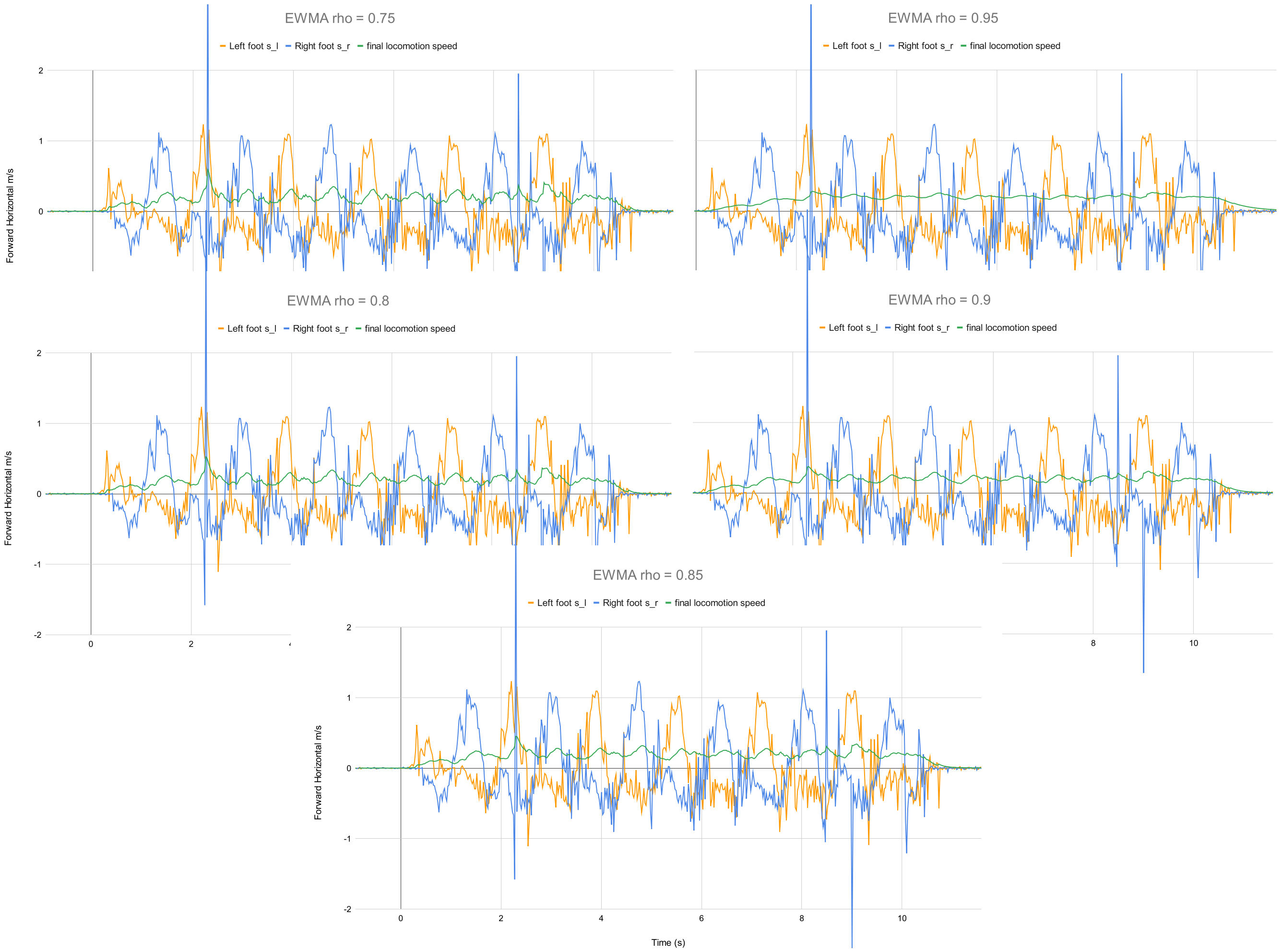


Fig. 15. EWMA tuning overview. Final chosen EWMA $\rho$ = 0.85

| | EWMA $\rho$ | | |
|---|---|---|---|
| | 0.85 | 0.1 | 0.95 |
| Average left foot clamped $>$ 0 (left lifted foot) | 0.1493 | 0.1493 | 0.1493 |
| Average right foot clamped $>$ 0 (right lifted foot) | 0.1493 | 0.1493 | 0.1493 |
| Average right foot clamped $<$ 0 (right standing foot) | -0.1448 | -0.1448 | -0.1448 |
| Average right foot not clamped (should be close to 0) | 0.0046 | 0.0046 | 0.0046 |
| Average user speed (should be close to 0) | 0.0067 | 0.0067 | 0.0067 |
| Average EWMA clamped $>$ 0 (lifted foot) | 0.1493 | 0.1493 | 0.1490 |
| Average EWMA clamped $<$ 0 (standing foot) | -0.1447 | -0.1448 | -0.1439 |
| Locomotion speed output both feet | 0.1618 | 0.1619 | 0.1612 |
| Speed scale ratio | 1.0838 | 1.0841 | 1.0797 |

TABLE III

Averages of speeds for the chosen $\rho$ (0.85), and two extremes (0.1 and 0.95). The averages displayed were used during steps in the locomotion speed calculation, or to check values to provide logical outputs. This was done with the same data used to design the locomotion speed while the user walked straight on the same spot in the room. The speed scale ratio was calculated by taking the locomotion speed output and dividing it by the average left foot clamped $>$ 0 (left lifted average foot speed).

located, the hip aligns, followed by the feet. On the target, they stand still and repeat. At the end of the trial, the user is walking the curved path. Here, looking at steadily changing angles (plot going up or down), we can see the directions are mostly aligned. Only when starting to change curvature (assigning a new goal) when turning in the curves, we can see the head, hip, and feet directions splitting from each other, with the head taking the lead, followed by the hip and then feet.

### *I. Implementation in Unity*

The algorithm designed above was then implemented in Unity. There were some minor changes introduced here.

To also make the algorithm work with for example backward or sideways walking if the standing foot direction was used, instead of distinguishing lifted/standing foot with a positive/negative speed, we decided to distinguish the lifted and standing

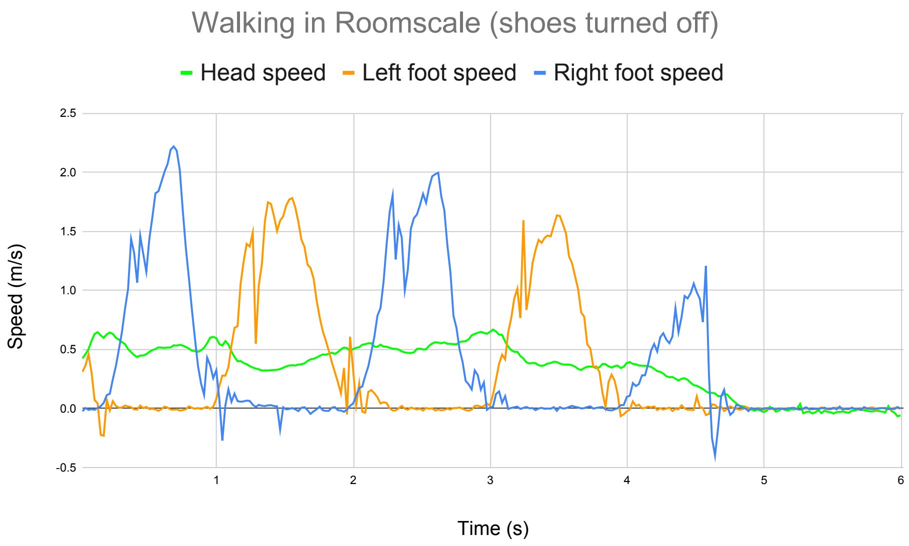


Fig. 16. The horizontal speeds are plotted along the axis the user was walking in a straight line at roomscale while the VR shoes are turned off (normal walking). We use this to compare our obtained locomotion speed including head motion, see Figure 17.

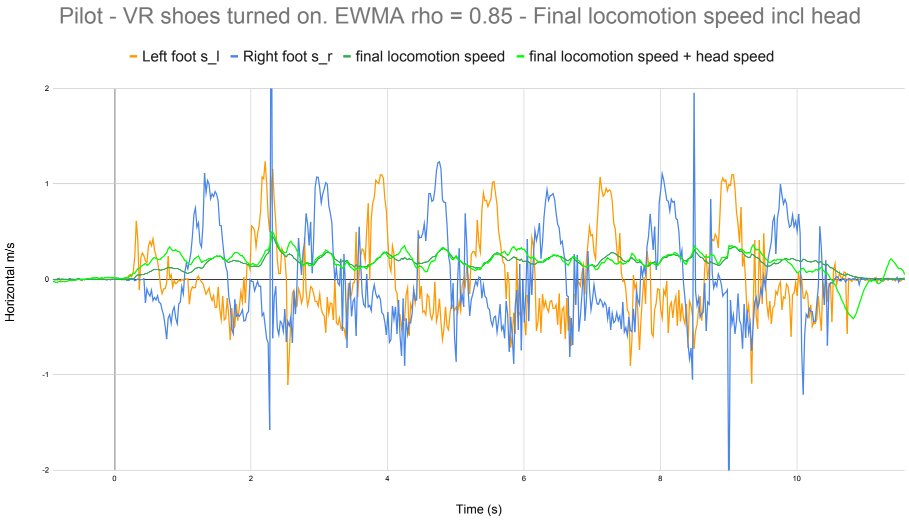


Fig. 17. The final locomotion speed obtained in the design process: positive speed EWMA (lifted foot), negative speed EWMA (standing foot), then averaging those, see Figure 14. We also include the final locomotion speed including head speed (VR headset resting on forehead) to compare against normal walking, see Figure 16.

foot based on height thresholds instead. This was tuned while walking in Unity, while live the color of the left/right foot changed based on whether it was lifted or standing, and tested while walking on the VR shoes visually, also see subsubsection II-B3.

We continued testing with the standing foot velocity, since in theory, this should be stationary with respect to the virtual floor. However, we noticed this was not the case: it skidded over it like the floor (locomotion speed) was moving too slowly. At the time, we thought this mismatch might come from multiple things: First, there is an offset from the tracker with respect to the toe, see Figure 19. However, that toe velocity is the velocity we desire since that is being driven across the ground for the back foot. This tracker offset causes a smaller horizontal component to be measured by the tracker. Additionally, the raw tracker velocity was used for the standing foot direction, which meant there was also noise in there, which was not noticeable in the optical flow smoothness in the headset but could be seen in the object indicating rotating to indicate the direction per frame. Applying this direction with noise and integrating it with the speed could result in a "zig-zag" integration of the speed, resulting in a slower final speed, see Figure 20. We applied a simple correction since our focus was getting a speed similar to normal walking, which did not need to match perfectly, but only within reason to be unnoticeable by participants. We scaled up the speed until the standing foot matched with the virtual floor while playing tracker data from the pilot (factor 1.2).

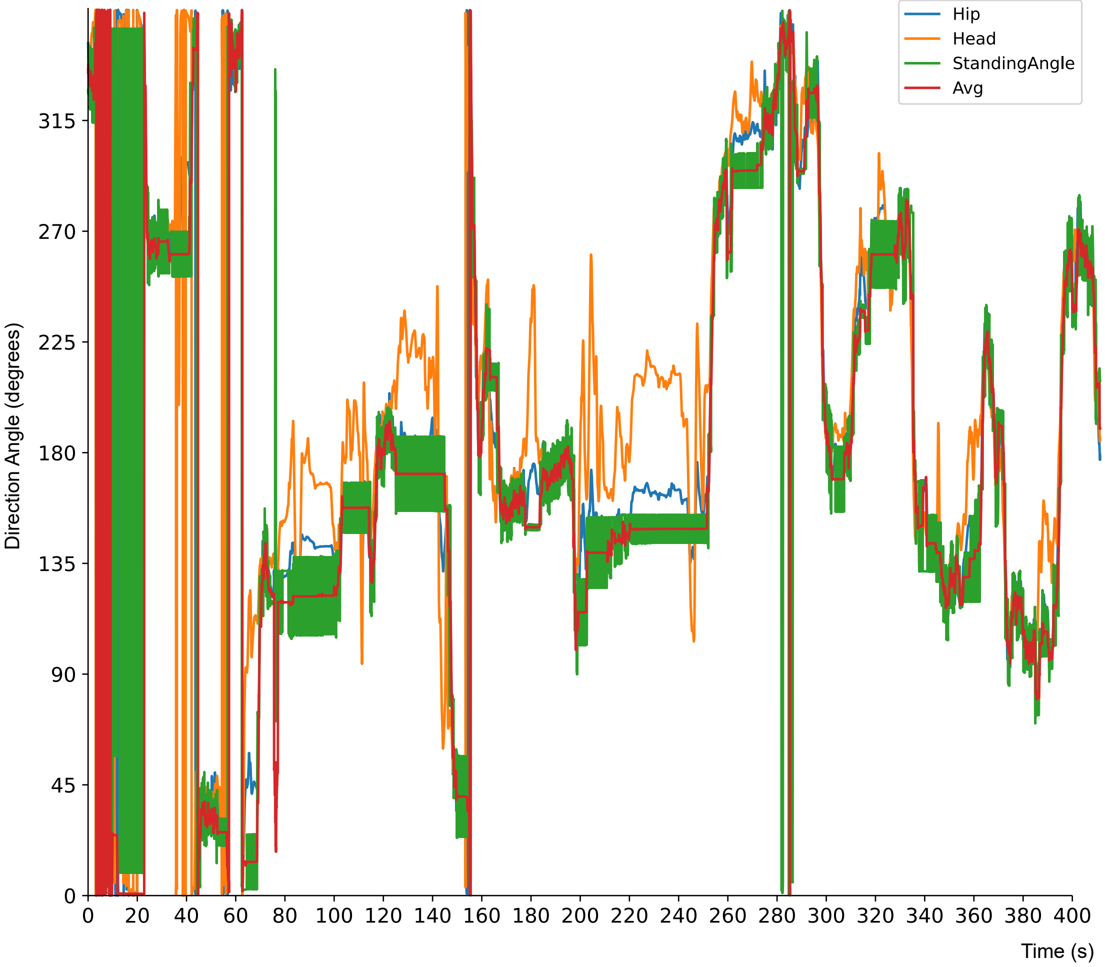


Fig. 18. The calculated (horizontal plane) angles of the directions $d$ with respect to the world x-axis for a user using the standing foot direction.

### J. Discussion data user test

The user test results on speed are given in subsection III-E, and is discussed in subsection IV-E. The results of the user test, showed the speed was thought to be slower than normal walking. It could be subjective only, but after analyzing the data from the user test, we believe that the scaling fix might not have been enough, and another component was influencing the speed objectively. Namely, for the implementation in Unity, we switched to detecting the lifted/standing foot based on a height threshold instead of the positive/negative speeds to detect these in our design phase. Since the speed was adjusted for the lifted foot in our calculations, if the lifted foot was not detected in all cases, it could have reduced the final speed. We noticed this seemed to be the case while plotting the lifted boolean for our user test data, see Figure 21. In hindsight, it is logical that using the height for detecting a lifted foot was not sufficient. The foot position during normal walking describes the path shown in Figure 22. This means that as the foot is still in the air, the heel dips and is near the same as the standing foot tracker height, while it should be considered a lifted foot. Therefore we believe this oversight is one of the main reasons, other than discussed in subsection A-I, our user-tested speed did not result in the 1:1 speed expected.

Additionally, we revisited our speed design algorithm. Here, we noticed that the speed, although 1:1 in our table, was relatively low: only 0.15 m/s. Therefore, we attempted to find out why this was since the shoes were slow, but we expected not 15 cm/s slow. Therefore we calculated the average based on the minimum between the two shoe speeds: taking the shoe that went the fastest backward at all times, see Figure 23. This resulted in a speed of 0.26 m/s. 26 cm/s still is very slow compared to a normal walking speed (1.5 m/s was found for casual walking at 90% confidence interval [82]). We think this difference is because we considered a single axis of movement in the direction the user was facing in the pilot (not using the total horizontal velocity). Therefore part of the velocity was not used in the design of the speed calculation with the pilot data. This is noticeable since looking at Figure 4, the speed seems to be near 0.7 m/s.

If we compare the 0.26 m/s used as input for our pilot data speed, we can see that means our average calculations in Table III, resulted in a slower optical flow: 0.26 m/s (real standing foot speed in the pilot/algorithm design) / 0.16 (locomotion speed calculated) = 0.62 ratio. This is below the perception threshold of 0.86, meaning that this could have caused the lower virtual speed participants experienced.

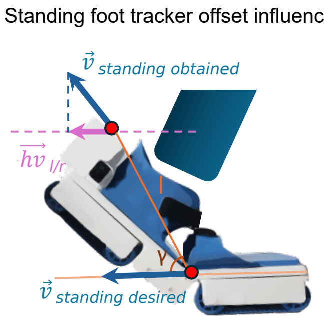


Fig. 19. The tracker has an offset l at the back of the foot with respect to the surface in contact with the ground. Since the tracker has this offset and the user lifts their heel during walking, a part of the horizontal velocity of the toe treadmill on the ground is not considered when using the data provided by the tracker.

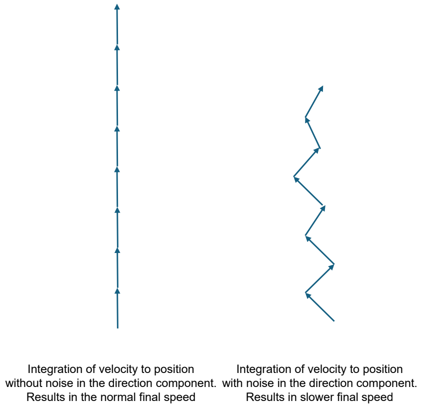


Fig. 20. If the direction component of the velocity vector contains noise, this could reduce the final overall speed.

### *K. Future work suggestions on speed*

To gain a better matching speed in future works, we suggest a few main corrections: lifted foot correction adjustment, noise direction correction, and tracker position adjustment.

*1) Correcting Lifted foot:* We calculated the lifted foot velocity by adding the standing foot velocity to it, so it could be averaged with the standing foot. Since the lifted foot did not seem to be detected well, see Figure 21.This can be resolved either of two ways: detect the lifted foot in a different way, or use an alternative calculation for the locomotion speed. The detection of the lifted foot could for example be performed with the same sensor on the bottom of the VR shoe that detects if the shoe is lifted for the VR shoe treadmill activation algorithm. Or, we could avoid using the lifted foot detection in our algorithm altogether, by changing how we combine the speeds. In hindsight, we believe a similar result could have been obtained by the following method. Since raw lifted foot signed speed (not containing standing foot correction), $s_{lifted_raw}$, on average is 0, we could simply add half of $s_{lifted_raw}$ to obtain a similar locomotion speed curve. This means we could just average the left and right shoe ($EWMA_{l/r}$) dividing only by 2, without needing any lifted foot detection. Alternatively, if backward or sideways walking is not required, one could also use the positive/negative speeds to distinguish the feet as was done in the design of the algorithm.

*2) Correcting speed standing foot:* We discovered that recalculating the locomotion ratio based on the minimum standing foot speed (VR shoe driving back the fastest), resulted in a 0.62 ratio, far below the detection threshold (0.85). Therefore, instead of calculating the speed with the average of the negative feet, we believe in future work this should be done with the standing foot moving the fastest, and still be combined with the lifted foot to cancel out the uneven locomotion speed of the standing foot, and reduce the latency, see subsection A-F and subsubsection A-K1.

*3) Correcting Tracker offset:* The offset of the tracker could have resulted in a smaller velocity than the user was actually walking, see Figure 19. In the future, one could attempt to correct this by using the angular velocity in combination with distance $l$ to estimate the velocity at the front of the foot.

*4) Correcting noisy direction integration:* For future work on matching the speed, it should be investigated how big the impact of this noise is on the integration (Figure 20), by performing the same integration both with the noise and smoothed direction. Smoothing any of the directions $d$ caused a noticeable delay in turning the head while using the raw data did not seem to give any visual problems. This could be resolved by either analyzing how big this difference is walking a straight

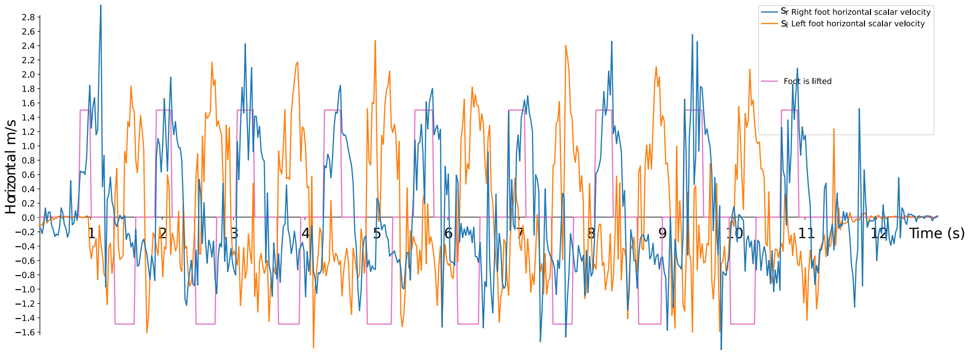


Fig. 21. Lifted foot labels on final user test data from Unity, positive in the plot: right foot lifted, negative in the plot: left foot lifted. It seems the lifted foot was considered not lifted anymore too early (as it was still moving forwards), which could have resulted in a slower final speed. Why the height threshold did not work properly can be seen in Figure 22.

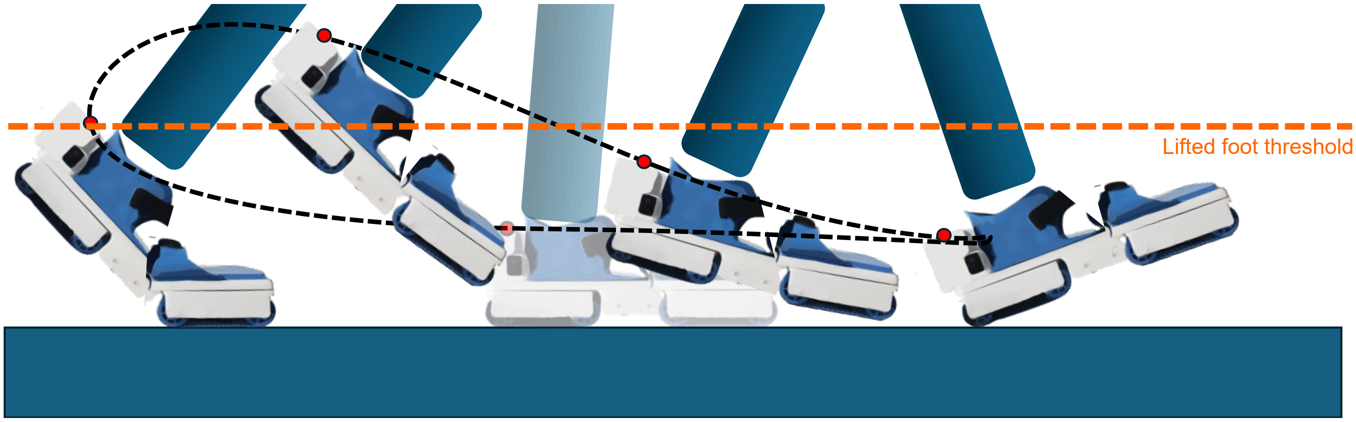


Fig. 22. This figure shows why using a height as the lifted foot threshold might not have worked as intended and cut off before the lifted foot was standing again, see Figure 21. This was plotted based off a regular foot trajectory. As mentioned in subsection IV-C, users had to plant down their feet more vertically using the VR shoes, so the curve might look different.

line, and then scaling the locomotion speed up to compensate for this difference precisely.

*5) Correcting the user "keeps moving after stopping":* A mention in the open question was that users kept moving after stopping. This could be due to tracker loss, but also due to the response time of the shoes. As seen in Figure 4, the VR speed stops near the same time the VR shoes stop driving backward. However, during the time the shoe is still driving backward, the user already stopped walking. In the future, it should be tested if this improves by setting the new $s_{l/r}$ to 0 (gradually or near instantly) if both feet are detected on the ground.

*6) Support for maneouvering:* Some people mentioned that they could not turn around without moving forward a bit. Therefore, a similar correction as in [74] should be possible to allow for maneouvering without triggering optical flow. They substracted a fixed speed from the locomotion speed, and clamped it. This increased the starting latency slightly, but shortened the stopping latency and allowed for maneouvering as long as the heels speeds stayed low. A similar trick could be applied here, although if a 1:1 speed is desired, this should be corrected once the locomotion is active.

### L. Discussion on previous works in transfer functions

The performance metrics that the reviewed papers mention, which are noted in Table II, are sometimes hard to compare between papers due to different and unclear definitions. Lee, Ahn, and Hwang [77] mentions a WIP step detection accuracy of 99.32%, and that it ignores most false positives. However, it is unclear how this has been calculated. The participants stopped 10 times in the experiment, and they mentioned the first step was never used for output. Each participant walked 1000 steps total. This means the error, excluding missed or extra recognitions, should already be over 1%. Therefore, it is not clear how they came to this accuracy, and whether it is comparable to the step detection accuracies from other works. It could be based on the percentage time with correct output, or these misses are simply not included in the accuracy calculation. In either case, this makes it hard to compare with other methods. Usoh, Arthur, Whitton, *et al.* [75] did not state the accuracy, but mentioned their errors improved compared to [76], who mentioned 91% accuracy. Therefore it is noted in the table to be better than 91% of the time in Table II. Tregillus and Folmer [78] also seemed to misinterpret "frames" to mean "steps" when they compared against [76]: "they did not start actual motion until 4 consecutive steps had been taken, which resulted in a large delay between

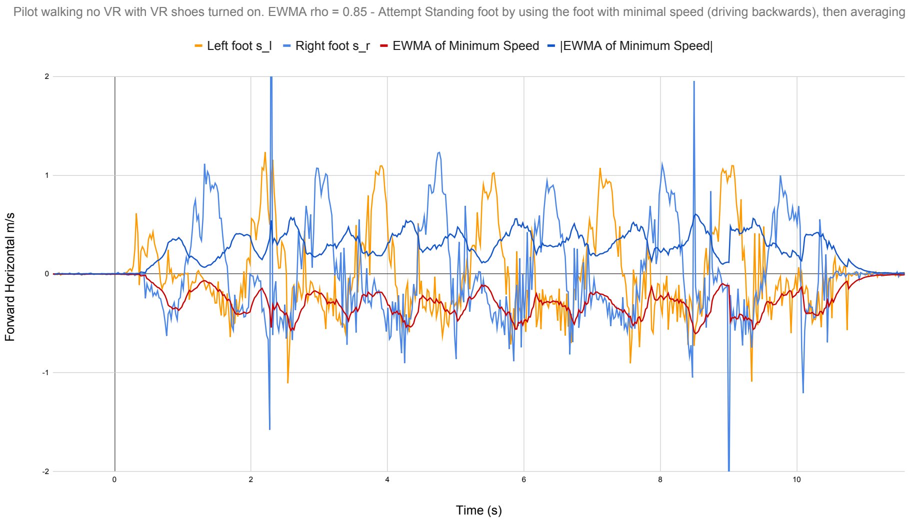


Fig. 23. This shows the algorithm based on the minimum of the two feet' speeds (since that foot drives back the fastest), filters this with EWMA, and then takes the absolute value to convert it to positive speed. This resulted in a non-smooth locomotion speed, although with a higher speed when this was averaged: 0.26 m/s. We believe this standing foot speed should be considered instead of the average clamped < 0 speed (see Figure 13) for future work.

walking and actually moving." This seems to be incorrect. Slater, Usoh, and Steed [76] mentions it takes 2 frames until the output is generated, so with 3 frames latency at 30 Hz, this is around 100 ms latency. It is also not clear if the neural network detects the data at the first step directly, or if there is a latency here as well. This is a generally observed problem: Many reviewed papers do not report or define the latency or accuracy of their method, while latency has proven to be important for precise navigation [74] and VR sickness [5].

The definitions of latency are also important to mention. For example, Feasel, Whitton, and Wendt [74] defines starting latency as the delay between the moment the foot leaves the ground and the moment virtual locomotion occurs and stopping latency as the moment the foot touches the ground and locomotion speed is zero. They also mention the following influences on the latency: the speed of the user (115ms vs 165ms for starting), filters, and offsets that eliminate false steps. Lee, Ahn, and Hwang [77] mention a latency of 44 ms, however, the virtual velocity only initiates after the second step, since the duration of the previous step is used to detect if a step is valid. Therefore the total latency is 1 to 2 steps + 44 ms. This makes it significantly longer than the 44 ms mentioned in their paper. Reporting accuracies and latencies should therefore be standardized and always reported to allow better comparisons between methods. There also was no standardized test for the user experience, so therefore comparing this between methods is also not possible. We did not measure accuracy (we did not perform step detection for which other works measured accuracy). However, we did measure latency. Since our lifted/standing foot implementation did not seem reliable and seems open to multiple interpretations, we therefore did not measure latency similar to Feasel, Whitton, and Wendt [74]. Our starting latency was defined between the point where the tracker registered horizontal motion to the point where the locomotion speed was at 1/4th of the average speed. Similarly, the stopping speed was the time at which both foot trackers had no horizontal movement, to the time the locomotion speed became zero. The 1/4th of the average locomotion speed is still vague and can be cheated by setting to locomotion speed to maximum/minimum instantly instead of slowly ramping it up or down, causing a jarring change in speed for the user. This is of course not desired, therefore latency should not be used as the sole measure of a locomotion algorithm. We provide all data and code to allow for future comparisons.

# APPENDIX B
## UX SCORES ANALYSIS

### A. Hypotheses

- Hypothesis 1: The general user experience score (all question groups weighted 1) is significantly different between all direction conditions (head, hip, average feet, standing foot).
- Hypothesis 2: Head will give the worst general user experience of all direction conditions.
- Hypothesis 3: Standing foot will give the best general user experience of all direction conditions.
- Hypothesis 4: Question groups Direction, Ease of Use, Perceived Errors, Satisfaction, and Control scores will differ the most between all direction conditions.
- Hypothesis 5: After inversion of the question scores so 100 is best and 1 is worst, all question group scores correlate positively with each other.
- Hypothesis 6: The higher the order number, the better the overall user experience.

### B. R command

Linear mixed model

```
ux_model = lme4::lmer( data = data_all, value
~ Condition * QuestionGroup + (1 |ParticipantID) + (1 |SubQuestion)
```

Hypothesis 1, 2, 3

```
conditionUXcomparison =
emmeans::emmeans(ux_model, list(pairwise ~ Condition), adjust="bonferroni")
```

Hypothesis 4

```
conditionQuestionGroupComparison = emmeans::emmeans(ux_model, ~ Condition| QuestionGroup)
pairs(conditionQuestionGroupComparison, simple = "Condition", adjust = "bonferroni")
```

Hypothesis 5

```
corTests = dataCorMat %>%
corrplot::cor.mtest(use = "pairwise.complete.obs")

corrplot::corrplot(corMat, p.mat = corTests$p, sig.level = c(0.001, 0.01, 0.05), insig =
    'label_sig', pch.cex = 0.9, order = "hclust" ) #,
```

Hypothesis 6, see section G.

### C. Result

Hypothesis 1, 2, 3

TABLE IV
PAIRWISE COMPARISON **UX SCORES DIRECTION CONDITIONS**

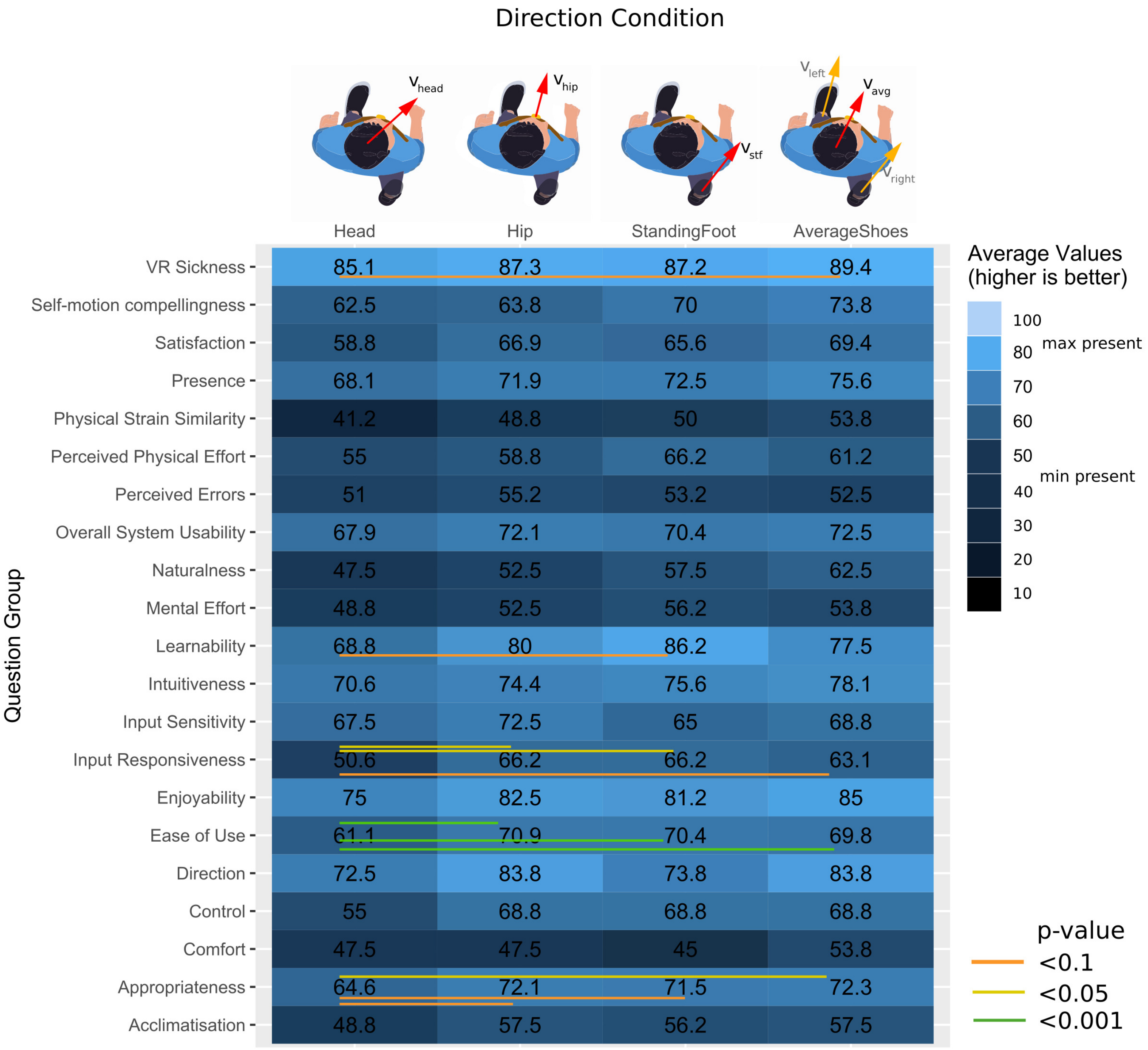
| contrast | estimate | std. error | p.value |
|---|---|---|---|
| Head - Hip | -6.5687956 | 1.289838 | 0.0000021177 |
| Head - StandingFoot | -6.7208648 | 1.289838 | 0.0000011292 |
| Head - AverageShoes | -8.3312673 | 1.289838 | 0.0000000006 |
| Hip - StandingFoot | -0.1520692 | 1.289833 | 1.0000000000 |
| Hip - AverageShoes | -1.7624717 | 1.289833 | 1.0000000000 |
| StandingFoot - AverageShoes | -1.6104025 | 1.289833 | 1.0000000000 |

Hypothesis 4:

The results are visually represented in Figure 24 in this appendix, and visually summarized in Figure 9, and Figure 8 in the main report. For a more detailed p-value table overview, see Appendix I.

Hypothesis 5: The results are visually represented in Figure 25.

Hypothesis 6: See section G.



| Question Group | Head | Hip | StandingFoot | AverageShoes |
|---|---|---|---|---|
| VR Sickness | 85.1 | 87.3 | 87.2 | 89.4 |
| Self-motion compellingness | 62.5 | 63.8 | 70 | 73.8 |
| Satisfaction | 58.8 | 66.9 | 65.6 | 69.4 |
| Presence | 68.1 | 71.9 | 72.5 | 75.6 |
| Physical Strain Similarity | 41.2 | 48.8 | 50 | 53.8 |
| Perceived Physical Effort | 55 | 58.8 | 66.2 | 61.2 |
| Perceived Errors | 51 | 55.2 | 53.2 | 52.5 |
| Overall System Usability | 67.9 | 72.1 | 70.4 | 72.5 |
| Naturalness | 47.5 | 52.5 | 57.5 | 62.5 |
| Mental Effort | 48.8 | 52.5 | 56.2 | 53.8 |
| Learnability | 68.8 | 80 | 86.2 | 77.5 |
| Intuitiveness | 70.6 | 74.4 | 75.6 | 78.1 |
| Input Sensitivity | 67.5 | 72.5 | 65 | 68.8 |
| Input Responsiveness | 50.6 | 66.2 | 66.2 | 63.1 |
| Enjoyability | 75 | 82.5 | 81.2 | 85 |
| Ease of Use | 61.1 | 70.9 | 70.4 | 69.8 |
| Direction | 72.5 | 83.8 | 73.8 | 83.8 |
| Control | 55 | 68.8 | 68.8 | 68.8 |
| Comfort | 47.5 | 47.5 | 45 | 53.8 |
| Appropriateness | 64.6 | 72.1 | 71.5 | 72.3 |
| Acclimatisation | 48.8 | 57.5 | 56.2 | 57.5 |

Fig. 24. An overview of the user experience scores for each question group and condition. The horizontal axis shows the four direction conditions tested. The vertical axis shows each question group. The scores are averages from the questions of the corresponding group and are flipped and scaled to match the 1 to 100 scores, with 100 being the best user experience (e.g. lowest VR sickness, highest naturalness). The conditions of avg feet, hip, and standing foot velocity have significantly higher scores than head-oriented walking ($p \leq 0.0005$) as indicated by the green brackets. The Ease of Use, Input Responsiveness and Appropriateness are the main contributors to this, as can be seen by the yellow ($p < 0.05$) and orange ($p < 0.1$) significance brackets. Aside from the indicated significant differences with the head-oriented condition comparison, no significant differences were found among the other conditions.

### D. Observations

- 1: As seen in Table IV, the general user experience score is only significantly different between the head direction and all other direction conditions ($p < 0.000003$).
- 2: The head gives the worst user experience.
- 3: The standing foot, hip, and average feet conditions do not significantly differ from each other ($p = 1$).
- 4: The largest mean difference between head and another condition with the largest difference, was Learnability (18 to standing foot, $p = 0.08$), Input Responsiveness (15.6 to standing foot $p = 0.01$), Naturalness (15 to average shoes $p = 0.2$), Physical Strain Similarity (12.6 to average shoes $p = 0.46$), Self-motion Compellingness (11.3 to average feet $p = 0.67$), Satisfaction (10.6 to average feet $p = 0.2$), Ease of Use (9.8 to hip $p = 0.0014$), Acclimatisation (8.7 to average shoes $p = 1$), and Appropriateness (7.7 to average shoes $p = 0.045$). VR sickness (4 to average shoes $p = 0.092$). Note that while Acclimatisation, Appropriateness, and VR sickness had much less difference between means than for example

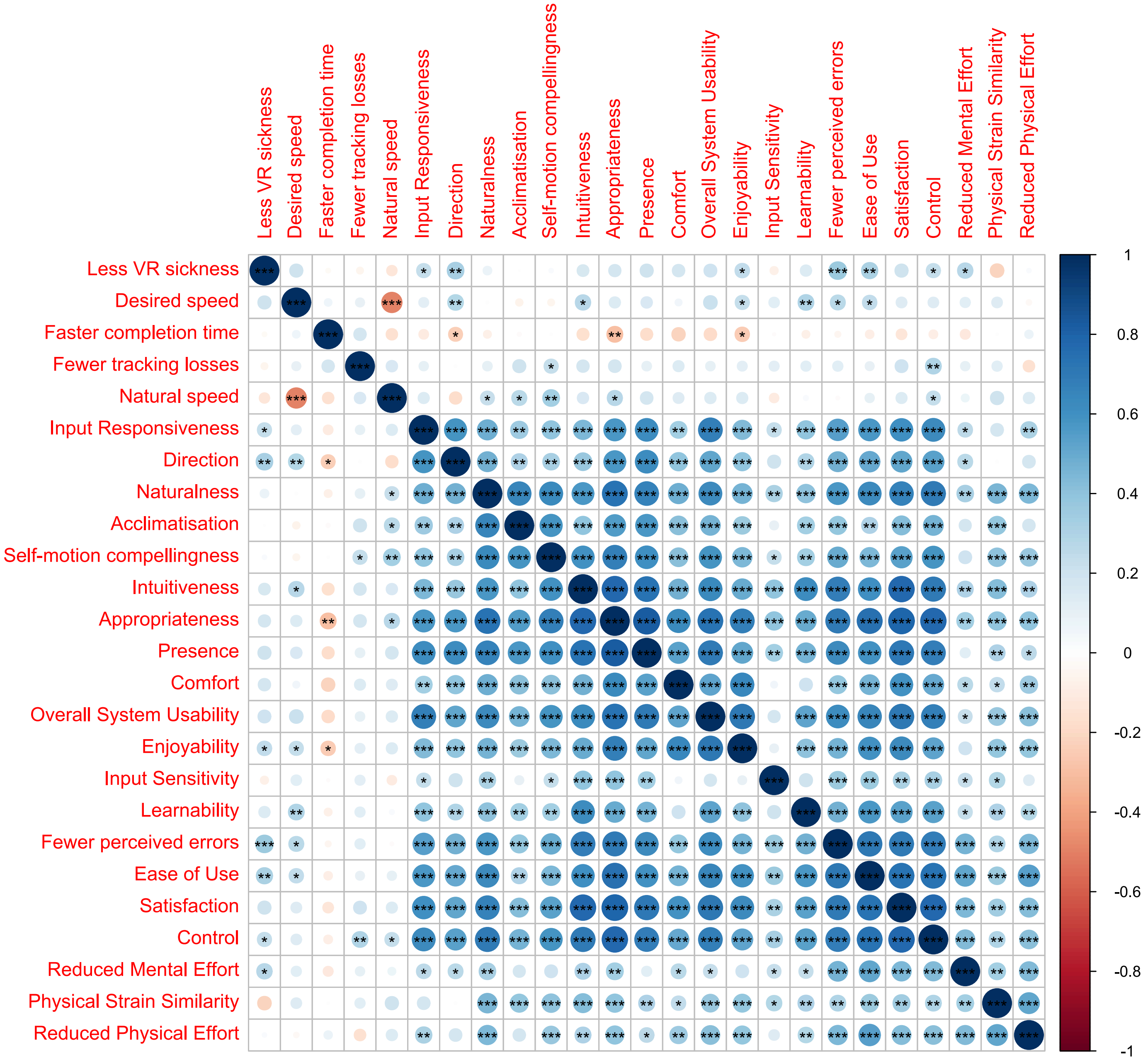


Fig. 25. This matrix shows the correlation between question groups after scaling and inversion so higher scores give a better user experience. Stars indicate p-value levels, with $*** < 0.001$, $** < 0.01$, and $* < 0.05$. Most user experiences positively correlate with each other. Only the speeds inversely correlate, but that is because due to their scale where the middle is best, these were not inverted, while the question was asked reverse. We mainly use this matrix to look at the effects of tracking loss and completion time, and to check if there are any other surprising results.

Naturalness, but were found more significant.

- 5: No question groups correlate significantly negatively with each other, and most question groups correlate significantly with each other. The only questionnaire-related question groups that correlated less with most of the other question groups were Input Sensitivity (not significantly correlated with 7 question groups), Mental Effort (not significantly correlated with 4 question groups), Perceived Physical Effort (not significantly correlated with three question groups), and Physical Strain Similarity (not significantly correlated with two question groups).
- 6, see section G.

### *E. Discussion*

The head was indeed experienced worse compared to the other conditions since participants could not look in a separate direction from their current movement direction. In normal walking, people first move their eyes towards the target, then the head orientation follows anticipating the future walking direction by 200 to 600ms before the body turns. However, the standing foot, average feet, and hip did not significantly differ from each other. One explanation is that these are less connected to the future direction anticipation like the head, and therefore less disturbed by our walking habit. Additionally, it is probably due to the directions between lower body directions not differing a lot from each other during the experiment. While participants could look in any direction while walking, the environment did not invite or require them to do so. Rotating the upper body especially, caused by aiming a weapon to the side while walking forward could cause these directions to differ more. Therefore, it could be these direction conditions do become significantly different when in an environment where this is required such as a shooting game with a walking goal and targets on the sides.

Note that Ease of Use, Acclimatisation, and Appropriateness had much less difference between means of different direction conditions than for example Naturalness, but Naturalness was not significantly different between conditions, while the others were. This is due to the number of questions asked: Ease of Use, Appropriateness, and Input Responsiveness had a larger number of questions (7, 6, and 2 questions respectively, versus only a single question for Naturalness. Therefore aside from naming the largest differences in means (Learnability, Input Responsiveness, and Naturalness), and most significantly different groups (Ease of Use, Acclimatisation, and Appropriateness), no clear conclusion can be drawn on the question groups that were most impacted by the direction condition.

Similarly, the single question explicitly asking if the direction matched expectations did not give a significantly different result between any direction conditions. However, 8/20 participants did actively notice differences in turning or direction and wrote this down in the open questions when asked about differences they noticed, although often it was a guess and participants were not fully sure. Especially while comparing the head condition people mentioned they had no good control over direction, or were moving with an offset angle.

All question groups were correlated positively with Ease of Use, which gave the concluding significant differences between direction conditions by itself. While Input Sensitivity, Mental Effort, Perceived Physical Effort, and Physical Strain Similarity did not significantly correlate with all other question groups, these also did not contribute to the conclusions comparing the direction conditions. Therefore in any future endeavors looking at direction conditions for VR shoes, asking for just Ease of Use on the questionnaire could be sufficient.

## APPENDIX C
## VIRTUAL SPEED ANALYSIS

### A. *Hypotheses*

- Hypothesis 1: The virtual speed between any direction conditions does not feel significantly different.
- Hypothesis 2: The speed set in the virtual environment feels natural.
- Hypothesis 3: Participants do not want to walk faster in the virtual environment than they are physically walking.
- Other: Quantitative analysis from the open questions.

### B. *R command*

Hypothesis 1

```
mymodel = lm(data = data,
          value ~ QuestionName * directionCondition)

my_model_emmeans = emmeans::emmeans(mymodel, ~ directionCondition | QuestionName)
result = pairs(my_model_emmeans, simple = "directionCondition")
```

Hypothesis 2, 3

```
   mean(filter(dataUX_pivot_Speed, QuestionName == "I think the virtual speed felt natural
       compared to normal walking")$value)

   mean(filter(dataUX_pivot_Speed, QuestionName == "I wanted to move through the virtual
       environment")$value)
```

### C. *Result*

Hypothesis 1

TABLE V
VIRTUAL SPEED COMPARISON BETWEEN DIRECTION CONDITIONS

| contrast | estimate | std. error | p.value |
|---|---|---|---|
| I think the virtual speed felt natural compared to normal walking (1 Too slow, 5 Too Fast) | | | |
| AverageShoes - Head | $4.00 \times 10^{-1}$ | $2.97 \times 10^{-1}$ | 0.5335 |
| AverageShoes - Hip | $3.50 \times 10^{-1}$ | $2.97 \times 10^{-1}$ | 0.6403 |
| AverageShoes - StandingFootVelocity | $1.14 \times 10^{-15}$ | $2.97 \times 10^{-1}$ | 1.0000 |
| Head - Hip | $-5.00 \times 10^{-2}$ | $2.97 \times 10^{-1}$ | 0.9983 |
| Head - StandingFootVelocity | $-4.00 \times 10^{-1}$ | $2.97 \times 10^{-1}$ | 0.5335 |
| Hip - StandingFootVelocity | $-3.50 \times 10^{-1}$ | $2.97 \times 10^{-1}$ | 0.6403 |
| I wanted to move through the virtual environment | | | |
| AverageShoes - Head | $1.50 \times 10^{-1}$ | $2.97 \times 10^{-1}$ | 0.9576 |
| AverageShoes - Hip | $-2.00 \times 10^{-1}$ | $2.97 \times 10^{-1}$ | 0.9067 |
| AverageShoes - StandingFootVelocity | $-8.44 \times 10^{-16}$ | $2.97 \times 10^{-1}$ | 1.0000 |
| Head - Hip | $-3.50 \times 10^{-1}$ | $2.97 \times 10^{-1}$ | 0.6403 |
| Head - StandingFootVelocity | $-1.50 \times 10^{-1}$ | $2.97 \times 10^{-1}$ | 0.9576 |
| Hip - StandingFootVelocity | $2.00 \times 10^{-1}$ | $2.97 \times 10^{-1}$ | 0.9067 |

Hypothesis 2, 3

- Mean conditions: too slow/fast compared to natural walking question = 2.6625
- Mean conditions: I wanted to go slower/faster question = 3.4625

### D. *Observations*

- There was no significant difference in speed between the different direction conditions ($p$ >0.53).
- People felt they were virtually walking slightly slower than their natural walking speed.
- People felt they wanted to move slightly faster through the virtual environment.

### *E. Discussion*

People felt they were moving slower than their natural walking speed. Some participants mentioned they found the speed in VR similar to what they were walking, but since they were limited by the VR shoe speed, it cost more energy, also see subsection IV-C. The following was described by participants in the open questions, sometimes contradicting each other with possible explanations from the researcher. It did feel like a natural speed, but the distances in VR felt longer and more tiring. The speed for the museum part was good, but when walking the long curved hallway it felt too slow, probably due to walking a long boring path, and possibly a lack of optical flow of the grey walls. However, another participant contradicted this: they felt the virtual environment did not result in the same speed as they were walking, and VR should move faster to feel like it matched. Some said they felt a lack of control over speed, which was probably due to tracker losses since this would cause them to suddenly shoot forward after recovering the tracker, or moving while they did not intend to, see subsection IV-G. Another mention was that they kept moving after stopping. This could be due to tracker loss, but also due to the response time of the shoes. As seen in Figure 4, the VR speed stops near the same time the VR shoes stop driving backward. In the future, it could be tested if this improves by setting the new $speed_l$ and $speed_r$ to 0 (gradually or instantly) if both feet are detected on the ground. Additionally, some people mentioned they could not turn around without moving forward a bit. Therefore an option to detect this kind of maneuvering motion should be implemented to prevent undesired movement.

In theory, the speed between different conditions should be the same, and as expected, the speed was not perceived as significantly different between direction conditions. However since there were only two questions on this which were processed separately, it could be people did notice differences, but it is not visible due to the limited sample size. Additionally, tracker losses could have caused more noise in participants' perception of speed.

# APPENDIX D
## VR SICKNESS ANALYSIS

### A. Hypotheses

- Hypothesis 1: VR sickness is significantly different between all direction conditions.
- Hypothesis 2: VR sickness increases with the order of the condition test within a participant.
- Hypothesis 3: VR sickness correlates with question groups Perceived Errors, Control, direction, Input Sensitivity, Input Responsiveness, and Comfort.

### B. R command

Hypothesis 1: See the R command of hypothesis 4 in Appendix B.
Hypothesis 2:

```
my_model = lmer(
  data = filter(data, QuestionGroup == "VR Sickness"),
  value ~ LocomotionTechnique * orderCondition +
    (1 | ParticipantID)
)
```

Hypothesis 3: See the R command of hypothesis 5 in Appendix B.

### C. Result

Hypothesis 1: See the result of hypothesis 4 VR sickness in Appendix B and p-value tables in Appendix I.
Hypothesis 2:

TABLE VI
VR SICKNESS ORDER EFFECT

| order number | estimate | std. error | p.value |
|---|---|---|---|
| orderCondition1 - (Intercept) | 80.4182 | 4.0467 | 0.0000 |
| orderCondition2 | 9.2524 | 5.7229 | 0.1156 |
| orderCondition3 | 8.2007 | 5.7269 | 0.1617 |
| orderCondition4 | 1.2470 | 3.4031 | 0.7141 |

Hypothesis 3: See the result of hypothesis 5 VR sickness in Appendix B.

### D. Observations

- VR sickness is not significantly different between conditions ($p > 0.092$).
- The VR sickness is not significantly different for the order of the condition test ($p > 0.11$).
- According to Figure 25, VR sickness correlates significantly with Perceived Errors ($p < 0.001$), Ease of Use ($p < 0.01$), direction ($p < 0.01$), enjoyability ($p < 0.05$), Control ($p < 0.05$), Input Responsiveness ($p < 0.05$), and Mental Effort ($p < 0.05$).

### E. Discussion

VR sickness is not significantly different between conditions. Additionally, the VR sickness is not significantly different for the order of the condition test, however, this could be due to the short time in VR of only on average 4 minutes per trial, with a break of around 10 minutes filling in the Questionnaire. In previous work, VR sickness seems to occur after the first 10 minutes [69]. While most UX factors are correlated, VR sickness seems to have relatively few UX factors it correlates with. VR sickness seems to correlate with Perceived Errors, Direction, Ease of Use, Input Responsiveness, Control, and Mental Effort, which are mostly similar to our expectations with the exceptions of Comfort, Input Sensitivity, and Mental Effort. It might be that this is due to the low VR sickness in general and correlates with more UX factors with higher VR sickness, for example induced in a longer experiment duration.

## APPENDIX E
## COMPLETION TIME ANALYSIS

The main logic of the hypothesis: More time spent to complete a test means more frustrating controls, and therefore a worse user experience. Additionally, the completion time is expected to become faster for a higher order number. The participant already walked through the experiment multiple times and has grown used to walking on shoes and the virtual environment.

- Hypothesis 1: More time spent is correlated with a worse user experience score.
- Hypothesis 2: The time spent is significantly different between all direction conditions.
- Hypothesis 3: The completion time is faster with increasing order number.
- Hypothesis 4: Tracker loss increases with faster completion time.

### A. *R command*

Hypothesis 1: See the R command of hypothesis 5 in Appendix B.

Hypothesis 2:

```
regTime <- lmer(data = dataTime,
          completionTime ~ LocomotionTechnique + (1 | ParticipantID))

regTime_emmeans = emmeans::emmeans(regTime, ~ LocomotionTechnique)
pairs(regTime_emmeans, adjust = "bonferroni")
# Bonferroni adjusted, we desire to find significance, so we have to be extra careful in
    concluding significance. Adjusting for multiple comparisons makes it less likely we find
    any significant results.
```

Hypothesis 3:

Hypothesis 4:

### B. *Result*

Hypothesis 1: The correlation matrix is shown in Figure 25.

Hypothesis 2

TABLE VII
P-VALUES FOR COMPARISON BETWEEN DIRECTION CONDITIONS ON COMPLETION TIME

| contrast | estimate | std. error | p.value |
|---|---|---|---|
| AverageShoes - Head | 7.902637 | 16.32685 | 1.0000000 |
| AverageShoes - Hip | -16.927761 | 16.32685 | 1.0000000 |
| AverageShoes - StandingFoot | -5.473662 | 16.07130 | 1.0000000 |
| Head - Hip | -24.830398 | 16.32685 | 0.8043422 |
| Head - StandingFoot | -13.376299 | 16.07130 | 1.0000000 |
| Hip - StandingFoot | 11.454099 | 16.07130 | 1.0000000 |

### C. *Observations*

- 1: A faster completion time had a negative impact on Appropriateness ($p < 0.01$), Direction ($p < 0.05$), and Enjoyability ($p < 0.05$).
- 2: There is no significant difference in completion time between the conditions ($p > 0.8$).
- 3:
- 4:

### D. *Discussion*

Interestingly, a faster completion time had a negative impact on some user experience question groups: Appropriateness, enjoyability, and direction. This could be explained by the participants attempting to rush through the experiment, and in the process having trouble with the VR shoes and keeping balance, or running into other control issues due to this. Especially

Appropriateness is about preferring this locomotion option for this task and being able to achieve what they wanted. From this, it is unclear whether being able to achieve the direction they desired is about the virtual direction condition, or having difficulty turning corners while walking on the VR shoes themselves. The time completion was not different between the conditions. This means participants were not able to walk faster in one condition than another, but only the user experience differed.

# APPENDIX F
# TRACKER LOSS ANALYSIS

## A. Hypotheses

- Hypothesis 1: More tracker losses had a negative impact on the user experience scores.
- Hypothesis 2: Tracker loss was not significantly different between direction conditions.
- Hypothesis 3: The tracking loss does not vary with the order number.

## B. R command

Hypothesis 1: See the R command of hypothesis 5 in Appendix B.
Hypothesis 2:

```
reg1 <- lmer(data = data_all,
         NrTrackerlosses ~ LocomotionTechnique + (1 | ParticipantID))

my_model_emmeans = emmeans::emmeans(reg1, ~ LocomotionTechnique)
pairs(my_model_emmeans, adjust = "none")

#not adjusted for multiple comparisons. We desire to find insignificance, so we have to be
    extra strict on concluding any insignificance. If this would be adjusted, we would find
    insignificant results more easily.
```

Hypothesis 3: see section G.

## C. Result

Hypothesis 1: The correlation matrix with user experience scores is given in Figure 25.
Hypothesis 2:

TABLE VIII
PAIRWISE COMPARISON **TRACKING LOSS LOCOMOTION TECHNIQUES**

| contrast | estimate | std. error | p.value |
|---|---|---|---|
| AverageShoes - Head | -1.031762117 | 1.923879 | 0.5939 |
| AverageShoes - Hip | -0.007940529 | 1.923879 | 0.9967 |
| AverageShoes - StandingFootVelocity | -0.734896833 | 1.896476 | 0.6999 |
| Head - Hip | 1.023821588 | 1.923879 | 0.5967 |
| Head - StandingFootVelocity | 0.296865284 | 1.896476 | 0.8762 |
| Hip - StandingFootVelocity | -0.726956304 | 1.896476 | 0.7030 |

Hypothesis 3: see section G.

## D. Observations

- 1: As visible in Figure 25, more tracker losses only negatively impacted Control ($p < 0.01$) and Self-motion Compellingness ($p < 0.01$). There is no significant correlation with any of the other question groups.
- 2: As visible in Table VIII, the tracker loss was not significantly different between conditions ($p > 0.59$).
- 3: see section G.

## E. Discussion

The trials lasted between two and a half minutes, and seven minutes (four minutes on average), Although the tracker losses were often mentioned in the open questions of the questionnaire, they seemed to have had little impact on the user experience scores. First of all, the number of tracker losses was not significantly different between conditions, so it does not have a final impact on results comparing between the conditions. When looking at the impact on user experience groups overall, only Control and Self-motion Compellingness seemed to have been impacted. Interestingly Perceived Errors, speed, and other related groups did not seem to be impacted. The result of a tracking loss was that the participant would not move during the trackerloss, but would suddenly shoot forwards with the missed speed between the previous and new tracking position directly after tracking recovery. When a participant asked if he should include the tracker loss experience to fill in the questionnaire, participants were told to try and ignore the instances of tracker loss. Since it is not especially noticeable in the UX data, but was often mentioned, this request seems to have reduced tracker loss impact in the questionnaire results.

# APPENDIX G
# ORDER EFFECTS ANALYSIS

## A. *Hypotheses*

- Hypothesis 1: The higher the order number, the better the overall user experience.
- Hypothesis 2: The higher the order number, the more negative VR sickness scores.
- Hypothesis 3: The higher the order number, the faster the completion time.
- Hypothesis 4: The higher the order number, the slower the virtual speed felt, and the faster they wanted to go. Explanation: both due to them getting used to the shoes and the environment and wanting to complete it faster, but also because the shoes started moving slightly slower as the battery drained.
- Hypothesis 5: The order did not influence the number of tracking losses.

## B. *R command*

Hypothesis 1:

```
my_model = lmer(
  data = dataUX_pivot,
  value ~ LocomotionTechnique * QuestionGroup + orderCondition +
    (1 | ParticipantID) + (1 | QuestionName)
)
summary(my_model)
```

Hypothesis 2:

```
my_model_emmeans = emmeans::emmeans(my_model, ~ orderCondition | QuestionGroup)
result = pairs(my_model_emmeans, adjust = "bonferroni")
print(result)
```

Hypothesis 3:

```
my_model = lmer(
  data = data_all,
  completionTime ~ LocomotionTechnique + orderCondition + (1 | ParticipantID)
)

my_model_emmeans = emmeans::emmeans(my_model, ~ orderCondition)
result = pairs(my_model_emmeans, simple = "orderCondition", adjust = "bonferroni")
print(result)
```

```
data_all %>%
  mutate(ParticipantID = factor(ParticipantID),
         orderCondition = as.numeric(orderCondition)) %>%
  ggplot() +
  geom_line(aes(x = orderCondition, y = completionTime)) +
  facet_wrap(vars(ParticipantID))
```

Hypothesis 4:

```
reg1 <- lmer(data = data_all,
             `I think the virtual speed felt natural compared to normal walking`
             ~ orderCondition + (1 | ParticipantID))

my_model_emmeans = emmeans::emmeans(reg1, ~ orderCondition)
result = pairs(my_model_emmeans, simple = "orderCondition", adjust = "none")
print(result)
```

```
reg1 <- lmer(data = data_all,
             `I wanted to move through the virtual environment`
             ~ orderCondition + (1 | ParticipantID))
summary(reg1)

my_model_emmeans = emmeans::emmeans(reg1, ~ orderCondition)
result = pairs(my_model_emmeans, simple = "orderCondition", adjust = "none")
print(result)
```

Hypothesis 5:

```
my_model_emmeans = emmeans::emmeans(reg1, ~ orderCondition)
result = pairs(my_model_emmeans, simple = "orderCondition", adjust = "bonferroni")
print(result)
```

### C. Result

Hypothesis 1: See Table IX.

TABLE IX
UX COMPARED BASED ON ORDER NUMBER

| contrast | | | estimate | SE | df | z.ratio | p.value |
|---|---|---|---|---|---|---|---|
| orderCondition1 | - | orderCondition2 | -4.654 | 1.3 | Inf | -3.587 | 0.002 |
| orderCondition1 | - | orderCondition3 | -5.586 | 1.3 | Inf | -4.305 | 0.0001 |
| orderCondition1 | - | orderCondition4 | -4.757 | 1.3 | Inf | -3.667 | 0.0015 |
| orderCondition2 | - | orderCondition3 | -0.933 | 1.3 | Inf | -0.719 | 1 |
| orderCondition2 | - | orderCondition4 | -0.104 | 1.3 | Inf | -0.08 | 1 |
| orderCondition3 | - | orderCondition4 | 0.829 | 1.3 | Inf | 0.639 | 1 |

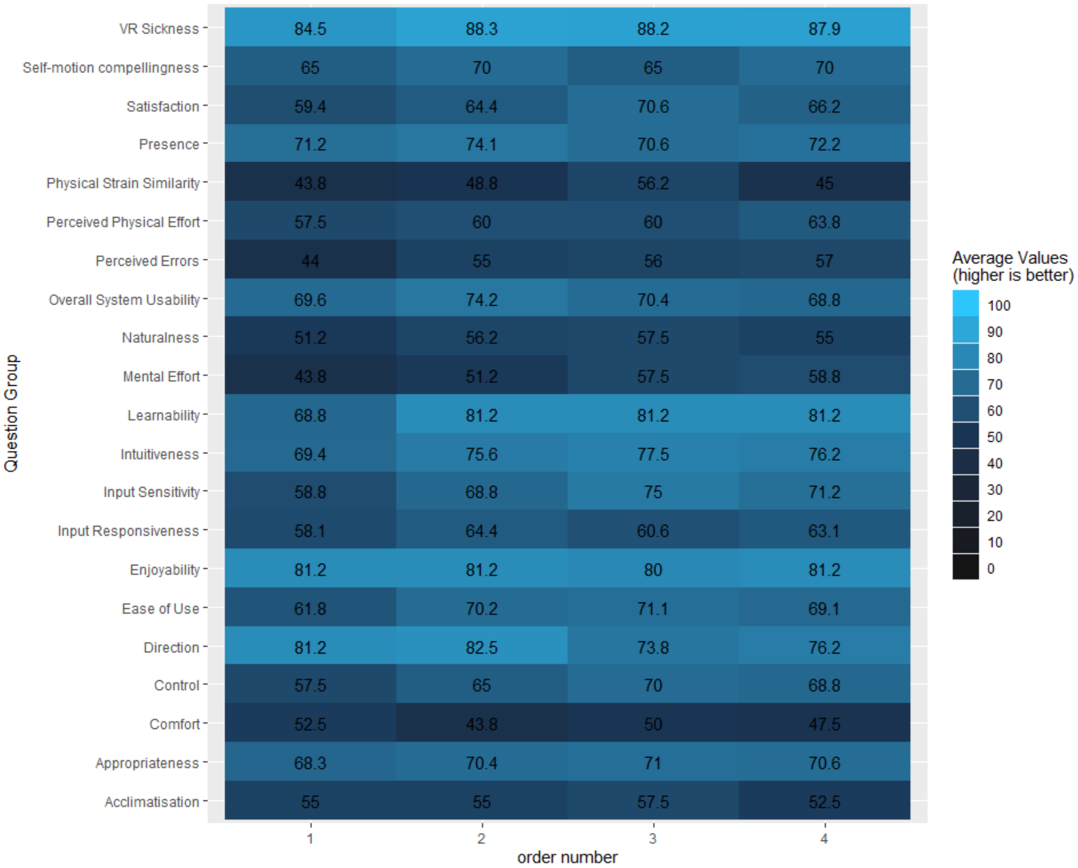


Fig. 26. Matrix comparing the user experience for each factor based on the order number. Overall the first order seems to be darker (worse), which is also what we see in Table IX.

Hypothesis 2: See Figure 27 and Table VI.
Hypothesis 3: See Table X Figure 28.
Hypothesis 4: See Table XI and Table XII.
Hypothesis 5: See Table XIII.

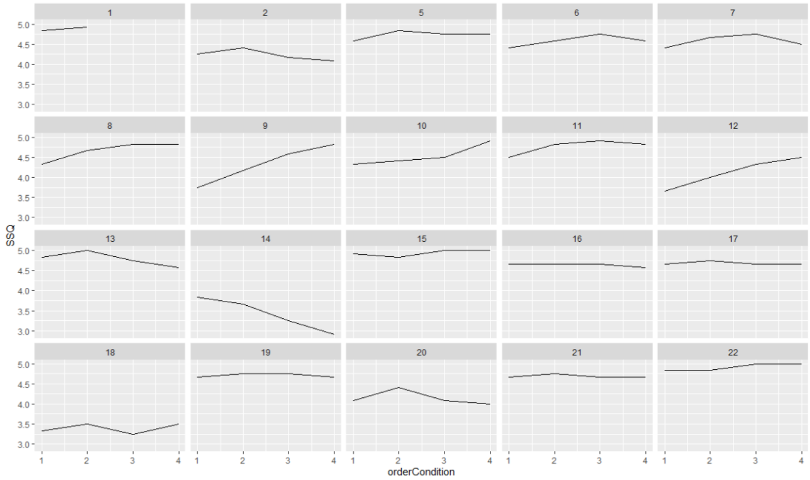


Fig. 27. For some users the VR sickness seemed to go up, for others down. Looking at the estimates, overall VR sickness went slightly down, but not significantly.

TABLE X
COMPLETION TIME COMPARED BETWEEN ORDER NUMBERS

| Completion time | | | | | | | |
|---|---|---|---|---|---|---|---|
| contrast | | | estimate | SE | df | t.ratio | p.value |
| orderCondition1 | - | orderCondition2 | 40 | 14.1 | 51.3 | 2.844 | 0.0383 |
| orderCondition1 | - | orderCondition3 | 52.2 | 14.3 | 51.6 | 3.648 | 0.0037 |
| orderCondition1 | - | orderCondition4 | 54.6 | 13.8 | 51 | 3.95 | 0.0014 |
| orderCondition2 | - | orderCondition3 | 12.2 | 14.6 | 52 | 0.839 | 1 |
| orderCondition2 | - | orderCondition4 | 14.6 | 14.1 | 51.3 | 1.039 | 1 |
| orderCondition3 | - | orderCondition4 | 2.4 | 14.3 | 51.6 | 0.167 | 1 |

### D. *Observations*

- Hypothesis 1: Order number 2, 3, and 4 all have a better overall user experience than order number 1. Especially perceived errors was different for the first trial compared to all the others $p < 0.003$.
- Hypothesis 2: The VR sickness was not significantly different between order numbers. For some participants, it seemed to increase, however overall it seemed to decrease (although not significantly). Participant 14 especially seemed to get progressively more VR sick with each trial. For most other participants it stayed relatively steady or even improved.
- Hypothesis 3: The first trial has a significantly higher completion time that all the other order nrs $p < 0.04$. This might

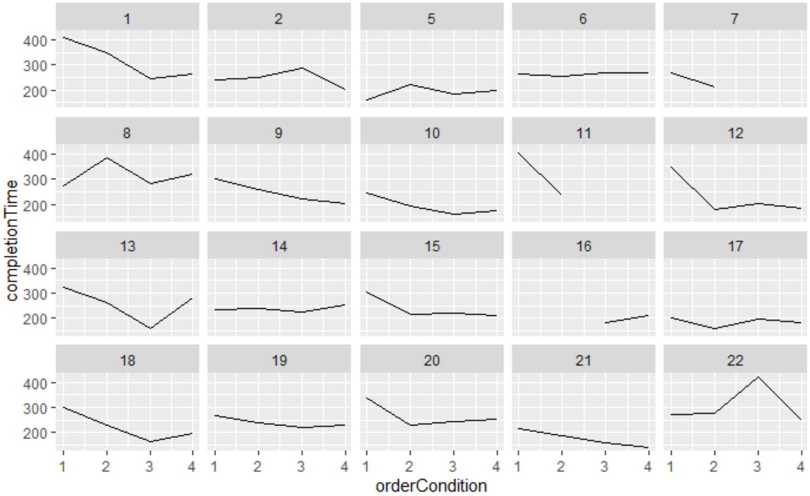


Fig. 28. Completion time shown for each participant over the order number

TABLE XI
QUESTION "VIRTUAL SPEED FELT NATURAL" COMPARED BETWEEN ORDER NUMBERS

| Speed felt natural | | | | | | | |
|---|---|---|---|---|---|---|---|
| contrast | | | estimate | SE | df | t.ratio | p.value |
| orderCondition1 | - | orderCondition2 | 0.25 | 0.274 | 57 | 0.913 | 0.3649 |
| orderCondition1 | - | orderCondition3 | 0.45 | 0.274 | 57 | 1.644 | 0.1057 |
| orderCondition1 | - | orderCondition4 | 0.25 | 0.274 | 57 | 0.913 | 0.3649 |
| orderCondition2 | - | orderCondition3 | 0.2 | 0.274 | 57 | 0.731 | 0.468 |
| orderCondition2 | - | orderCondition4 | 0 | 0.274 | 57 | 0 | 1 |
| orderCondition3 | - | orderCondition4 | -0.2 | 0.274 | 57 | -0.731 | 0.468 |

TABLE XII
QUESTION "WANTED TO GO FASTER" COMPARED BETWEEN ORDER NUMBERS.

| Wanted to go faster | | | estimate | SE | df | t.ratio | p.value |
|---|---|---|---|---|---|---|---|
| orderCondition1 | - | orderCondition2 | -0.3 | 0.224 | 57 | -1.341 | 0.1851 |
| orderCondition1 | - | orderCondition3 | -0.3 | 0.224 | 57 | -1.341 | 0.1851 |
| orderCondition1 | - | orderCondition4 | -0.25 | 0.224 | 57 | -1.118 | 0.2683 |
| orderCondition2 | - | orderCondition3 | 0 | 0.224 | 57 | 0 | 1 |
| orderCondition2 | - | orderCondition4 | 0.05 | 0.224 | 57 | 0.224 | 0.8239 |
| orderCondition3 | - | orderCondition4 | 0.05 | 0.224 | 57 | 0.224 | 0.8239 |

indicate users were still learning to walk on the shoes. Most participants continued to get faster for all trials, although this was mostly flattened after the first trial.

- Hypothesis 4: There was no significant difference between order numbers for speed that felt natural or wanting to go faster.
- Hypothesis 5: Tracker loss was not significantly different between order numbers.

### *E. Discussion*

We expected the user experience to be better with a higher order number. Indeed, order numbers 2, 3, and 4 all have a better user experience than trial number 1. Therefore we expect users were still learning to walk on the shoe, which negatively impacted their user experience. This is especially visible since Perceived Errors was an individual UX factor that was significantly different for the first trial compared to all others ($p < 0.003$).

VR sickness was not significantly different between order numbers. This was surprising since we expected VR sickness to increase over time, since users were exposed longer to VR. For most participants, the VR sickness instead flatlined or decreased with higher trials, although not significant. A single participant, 14, who had the highest SSQ score and was the only participant with no VR experience, showed that it indeed gradually worsened over time. Therefore we believe this should be retested with participants with less VR experience.

We expected the participants to be able to complete the trials faster as they gained more experience. This was indeed the case, but only for the first to second trial, indicating their learning curve was mostly leveled out after training one full run in VR.

We expected users to get used to the virtual speed and want to go faster for a higher order number. This was not the case ($p > 0.1$). However, since this is an analysis of a single question for which we have seen our sample size is not big enough to obtain significance, it could be this was affected but is not visible due to the statistical power.

As expected, the tracker losses also did not significantly differ with the order number.

TABLE XIII
TRACKER LOSS COMPARED BETWEEN ORDER NUMBERS.

| Tracker loss | | | estimate | SE | df | t.ratio | p.value |
|---|---|---|---|---|---|---|---|
| orderCondition1 | - | orderCondition2 | 2.106 | 1.87 | 51.7 | 1.124 | 1 |
| orderCondition1 | - | orderCondition3 | 3.932 | 1.9 | 52.3 | 2.065 | 0.2635 |
| orderCondition1 | - | orderCondition4 | 1.95 | 1.85 | 51.1 | 1.057 | 1 |
| orderCondition2 | - | orderCondition3 | 1.826 | 1.93 | 53 | 0.944 | 1 |
| orderCondition2 | - | orderCondition4 | -0.156 | 1.87 | 51.7 | -0.083 | 1 |
| orderCondition3 | - | orderCondition4 | -1.982 | 1.9 | 52.3 | -1.041 | 1 |

## APPENDIX H
## UX QUESTIONNAIRE

TABLE XIV: Questionnaire per question group and number indicating if the score was inverted to get a higher score for a better user experience. A minus means an inverted score, while the number indicates the type of scaling to get scores from 0 to 100 (see the bottom of the table).

| **Question description** | **Score Inverted** |
|---|---|
| VR sickness | |
| *Rate how much each symptom below is affecting you right now in a scale: 0 (None), 1 (Slight), 2 (Moderate), and 3 (Severe)* | |
| General discomfort | -2 |
| Fatigue | -2 |
| Headache | -2 |
| Eye strain | -2 |
| Difficulty focusing | -2 |
| Salivation increasing | -2 |
| Sweating | -2 |
| Nausea | -2 |
| Difficulty concentrating | -2 |
| Fullness of the Head | -2 |
| Blurred vision | -2 |
| Dizziness with eyes open | -2 |
| Dizziness with eyes closed | -2 |
| Vertigo | -2 |
| Stomach awareness | -2 |
| Burping | -2 |
| *(1 Strongly Disagree, 5 Strongly Agree unless otherwise specified)* | |
| *Ease of Use* | |
| I found the interface easy to use. | 1 |
| The interface was too complicated to use effectively. | -1 |
| I found it easy to move or reposition myself in the virtual environment. | 1 |
| I did not need any further help. | 1 |
| The interface interfered with the way I wanted to interact with the system. | -1 |
| I found it easy to undo mistakes and return to a previous state. | 1 |
| I was confused by the operation of the VR motion interface. | -1 |
| *Perceived Errors* | |
| The interfaces provided protection against trivial errors. | 1 |
| It was not possible to make silly mistakes. | 1 |
| The interface was very robust and reliable. | 1 |
| I kept making mistakes while interacting with the virtual environment. | -1 |

Continued on the next page ...

Continued from the previous page.

| **Question description** <br> *(1 Strongly Disagree, 5 Strongly Agree unless otherwise specified)* | **Score Inverted** |
|---|---|
| I was unaware of making mistakes. | -1 |
| *Appropriateness* | |
| The level of functionality control provided by the interface was appropriate for the task. | 1 |
| The functionality provided by the interface was ambiguous. | -1 |
| I would have preferred an alternative interface to carry out the task. | -1 |
| The interface was ideal for interacting with a virtual environment. | 1 |
| I had the right level of control over what I wanted to do. | 1 |
| I could not achieve what I wanted to do. | -1 |
| *Input Sensitivity* | |
| I found the interface too sensitive to use. | -1 |
| *Input Responsiveness* | |
| The response to user input was acceptable. | 1 |
| The response time did not affect my performance. | 1 |
| *Naturalness* | |
| Rate how natural you found the experience of walking and interacting in the virtual environment. (1 Very Low, 5 Very High) | 1 |
| *Physical Strain Similarity* | |
| Rate how different the physical strain was compared to normal walking. (1 Very different, 5 The same) | 1 |
| *Mental Effort* | |
| How mentally demanding was the task considering the interface used to perform it? (1 Not demanding, 5 Very demanding) | -1 |
| *Perceived Physical Effort* | |
| How physically demanding was the task considering the interface used to perform it? (1 Not demanding, 5 Very demanding) | -1 |
| *Satisfaction* | |
| The interface used to perform the specific task was satisfying. | 1 |
| The interface behaved in a manner that I expected. | 1 |
| *Self-motion compellingness* | |
| Rate whether you indeed felt as if you were moving while walking in the virtual environment overall. (1 Very Low, 5 Very High) | 1 |
| *Acclimatisation* | |
| Rate how quickly you forgot that you were not really walking. (1 Very slowly, 5 Very quickly) | 1 |
| *Presence* | |
| I got a sense of presence ie. of being there during the experience. | 1 |
| The behavior of the interface reduced my sense of presence. | -1 |
| I had a good sense of scales while moving and interacting with the virtual environment. | 1 |
| I often did not know where I was in the virtual environment. | -1 |
| *Overall System Usability* | |
| I thought that the interface worked against me. | -1 |
| The overall response time did not affect my performance. | 1 |
| I can see a real benefit in this style of manmachine interface. | 1 |
| *Enjoyability* | |

Continued on the next page ...

| Continued from the previous page. | |
|---|---|
| **Question description**<br>*(1 Strongly Disagree, 5 Strongly Agree unless otherwise specified)* | **Score Inverted** |
| I enjoyed carrying out the tasks. | 1 |
| *Control* | |
| I felt in control of the interface. | 1 |
| *Learnability* | |
| I found it difficult to learn how to use the interface. | -1 |
| Intuitiveness | |
| I did not have a clear idea of how to perform a particular function. | -1 |
| The interface did not work as expected. | -1 |
| *Comfort* | |
| I would be comfortable using this interface for long periods. | 1 |
| *Virtual speed* | |
| I think the virtual speed felt natural compared to normal walking (1 Much too slow, 5 Much too fast). | 0 |
| I wanted to move through the virtual environment (1 much slower, 5 much faster). | 0 |
| *Direction* | |
| I felt like I was moving in the direction I wanted to go to. | 1 |
| **Inversion definition:** | |
| -2 scaled from 0-3 to 0-100, and inverted | |
| 1 scaled from 1-5 to 0-100, and not inverted | |
| -1 scaled from 1-5 to 0-100, and inverted | |
| 0 analyzed separately, 3 out of 5 is best | |
| **Open Questions** | |
| *After each trial* | |
| Any comments, tips or remarks? | |
| *After all trials* | |
| Which conditions did you like or dislike? What was your favorite? | |
| Did you notice any differences between the conditions? If so, what did you notice? | |
| After VR shoe only training: Any comments, tips, suggestions? | |
| | End of the long table |

# APPENDIX I
## UX QUESTIONGROUP P-VALUES
### UX Hypothesis 4

TABLE XV: P-value table of comparisons between direction conditions per question groups.

| contrast | estimate | SE | p.value |
|---|---|---|---|
| Acclimatisation | | | |
| Head - Hip | −8.75 | 7.08 | 1.0000 |
| Head - StandingFoot | −7.50 | 7.08 | 1.0000 |
| Head - AverageShoes | −8.75 | 7.08 | 1.0000 |
| Hip - StandingFoot | 1.25 | 7.08 | 1.0000 |
| Hip - AverageShoes | $-8.72 \times 10^{-13}$ | 7.08 | 1.0000 |
| StandingFoot - AverageShoes | −1.25 | 7.08 | 1.0000 |
| Appropriateness | | | |
| Head - Hip | −7.50 | 2.89 | 0.0565 |
| Head - StandingFoot | −6.88 | 2.89 | 0.1039 |
| Head - AverageShoes | −7.71 | 2.89 | 0.0457 |
| Hip - StandingFoot | $6.25 \times 10^{-1}$ | 2.89 | 1.0000 |
| Hip - AverageShoes | $-2.08 \times 10^{-1}$ | 2.89 | 1.0000 |
| StandingFoot - AverageShoes | $-8.33 \times 10^{-1}$ | 2.89 | 1.0000 |
| Comfort | | | |
| Head - Hip | $5.33 \times 10^{-14}$ | 7.08 | 1.0000 |
| Head - StandingFoot | 2.50 | 7.08 | 1.0000 |
| Head - AverageShoes | −6.25 | 7.08 | 1.0000 |
| Hip - StandingFoot | 2.50 | 7.08 | 1.0000 |
| Hip - AverageShoes | −6.25 | 7.08 | 1.0000 |
| StandingFoot - AverageShoes | −8.75 | 7.08 | 1.0000 |
| Control | | | |
| Head - Hip | $-1.37 \times 10^{1}$ | 7.08 | 0.3119 |
| Head - StandingFoot | $-1.37 \times 10^{1}$ | 7.08 | 0.3119 |
| Head - AverageShoes | $-1.37 \times 10^{1}$ | 7.08 | 0.3119 |
| Hip - StandingFoot | $-1.78 \times 10^{-15}$ | 7.08 | 1.0000 |
| Hip - AverageShoes | 0.00 | 7.08 | 1.0000 |
| StandingFoot - AverageShoes | $1.78 \times 10^{-15}$ | 7.08 | 1.0000 |
| Direction | | | |
| Head - Hip | $-1.13 \times 10^{1}$ | 7.08 | 0.6710 |
| Head - StandingFoot | −1.25 | 7.08 | 1.0000 |
| Head - AverageShoes | $-1.13 \times 10^{1}$ | 7.08 | 0.6710 |
| Hip - StandingFoot | 1.00 | 7.08 | 0.9453 |
| Hip - AverageShoes | $-1.78 \times 10^{-15}$ | 7.08 | 1.0000 |
| StandingFoot - AverageShoes | $-1.00 \times 10^{1}$ | 7.08 | 0.9453 |
| Ease of Use | | | |
| Head - Hip | −9.82 | 2.67 | 0.0014 |
| Head - StandingFoot | −9.29 | 2.67 | 0.0031 |
| Head - AverageShoes | −8.75 | 2.67 | 0.0064 |
| Hip - StandingFoot | $5.36 \times 10^{-1}$ | 2.67 | 1.0000 |
| Hip - AverageShoes | 1.07 | 2.67 | 1.0000 |
| StandingFoot - AverageShoes | $5.36 \times 10^{-1}$ | 2.67 | 1.0000 |
| Enjoyability | | | |

Continued on the next page ...

Continued from the previous page.

| contrast | estimate | SE | p.value |
|---|---|---|---|
| Head - Hip | $-7.50$ | 7.08 | 1.0000 |
| Head - StandingFoot | $-6.25$ | 7.08 | 1.0000 |
| Head - AverageShoes | $-1.00 \times 10^{1}$ | 7.08 | 0.9453 |
| Hip - StandingFoot | 1.25 | 7.08 | 1.0000 |
| Hip - AverageShoes | $-2.50$ | 7.08 | 1.0000 |
| StandingFoot - AverageShoes | $-3.75$ | 7.08 | 1.0000 |
| Input Responsiveness | | | |
| Head - Hip | $-1.56 \times 10^{1}$ | 5.00 | 0.0107 |
| Head - StandingFoot | $-1.56 \times 10^{1}$ | 5.00 | 0.0107 |
| Head - AverageShoes | $-1.25 \times 10^{1}$ | 5.00 | 0.0748 |
| Hip - StandingFoot | $1.78 \times 10^{-15}$ | 5.00 | 1.0000 |
| Hip - AverageShoes | 3.13 | 5.00 | 1.0000 |
| StandingFoot - AverageShoes | 3.12 | 5.00 | 1.0000 |
| Input Sensitivity | | | |
| Head - Hip | $-5.00$ | 7.08 | 1.0000 |
| Head - StandingFoot | 2.50 | 7.08 | 1.0000 |
| Head - AverageShoes | $-1.25$ | 7.08 | 1.0000 |
| Hip - StandingFoot | 7.50 | 7.08 | 1.0000 |
| Hip - AverageShoes | 3.75 | 7.08 | 1.0000 |
| StandingFoot - AverageShoes | $-3.75$ | 7.08 | 1.0000 |
| Intuitiveness | | | |
| Head - Hip | $-3.75$ | 5.00 | 1.0000 |
| Head - StandingFoot | $-5.00$ | 5.00 | 1.0000 |
| Head - AverageShoes | $-7.50$ | 5.00 | 0.8031 |
| Hip - StandingFoot | $-1.25$ | 5.00 | 1.0000 |
| Hip - AverageShoes | $-3.75$ | 5.00 | 1.0000 |
| StandingFoot - AverageShoes | $-2.50$ | 5.00 | 1.0000 |
| Learnability | | | |
| Head - Hip | $-1.13 \times 10^{1}$ | 7.08 | 0.6710 |
| Head - StandingFoot | $-1.75 \times 10^{1}$ | 7.08 | 0.0803 |
| Head - AverageShoes | $-8.75$ | 7.08 | 1.0000 |
| Hip - StandingFoot | $-6.25$ | 7.08 | 1.0000 |
| Hip - AverageShoes | 2.50 | 7.08 | 1.0000 |
| StandingFoot - AverageShoes | 8.75 | 7.08 | 1.0000 |
| Mental Effort | | | |
| Head - Hip | $-3.75$ | 7.08 | 1.0000 |
| Head - StandingFoot | $-7.50$ | 7.08 | 1.0000 |
| Head - AverageShoes | $-5.00$ | 7.08 | 1.0000 |
| Hip - StandingFoot | $-3.75$ | 7.08 | 1.0000 |
| Hip - AverageShoes | $-1.25$ | 7.08 | 1.0000 |
| StandingFoot - AverageShoes | 2.50 | 7.08 | 1.0000 |
| Naturalness | | | |
| Head - Hip | $-5.00$ | 7.08 | 1.0000 |
| Head - StandingFoot | $-1.00 \times 10^{1}$ | 7.08 | 0.9453 |
| Head - AverageShoes | $-1.50 \times 10^{1}$ | 7.08 | 0.2040 |
| Hip - StandingFoot | $-5.00$ | 7.08 | 1.0000 |
| Hip - AverageShoes | $-1.00 \times 10^{1}$ | 7.08 | 0.9453 |
| StandingFoot - AverageShoes | $-5.00$ | 7.08 | 1.0000 |
| Overall System Usability | | | |

Continued on the next page …

Continued from the previous page.

| contrast | estimate | SE | p.value |
|---|---|---|---|
| Head - Hip | $-4.17$ | 4.09 | 1.0000 |
| Head - StandingFoot | $-2.50$ | 4.09 | 1.0000 |
| Head - AverageShoes | $-4.58$ | 4.09 | 1.0000 |
| Hip - StandingFoot | 1.67 | 4.09 | 1.0000 |
| Hip - AverageShoes | $-4.17 \times 10^{-1}$ | 4.09 | 1.0000 |
| StandingFoot - AverageShoes | $-2.08$ | 4.09 | 1.0000 |
| Perceived Errors | | | |
| Head - Hip | $-4.25$ | 3.16 | 1.0000 |
| Head - StandingFoot | $-2.25$ | 3.16 | 1.0000 |
| Head - AverageShoes | $-1.50$ | 3.16 | 1.0000 |
| Hip - StandingFoot | 2.00 | 3.16 | 1.0000 |
| Hip - AverageShoes | 2.75 | 3.16 | 1.0000 |
| StandingFoot - AverageShoes | $7.50 \times 10^{-1}$ | 3.16 | 1.0000 |
| Perceived Physical Effort | | | |
| Head - Hip | $-3.75$ | 7.08 | 1.0000 |
| Head - StandingFoot | $-1.13 \times 10^{1}$ | 7.08 | 0.6710 |
| Head - AverageShoes | $-6.25$ | 7.08 | 1.0000 |
| Hip - StandingFoot | $-7.50$ | 7.08 | 1.0000 |
| Hip - AverageShoes | $-2.50$ | 7.08 | 1.0000 |
| StandingFoot - AverageShoes | 5.00 | 7.08 | 1.0000 |
| Physical Strain Similarity | | | |
| Head - Hip | $-7.50$ | 7.08 | 1.0000 |
| Head - StandingFoot | $-8.75$ | 7.08 | 1.0000 |
| Head - AverageShoes | $-1.25 \times 10^{1}$ | 7.08 | 0.4637 |
| Hip - StandingFoot | $-1.25$ | 7.08 | 1.0000 |
| Hip - AverageShoes | $-5.00$ | 7.08 | 1.0000 |
| StandingFoot - AverageShoes | $-3.75$ | 7.08 | 1.0000 |
| Presence | | | |
| Head - Hip | $-3.75$ | 3.54 | 1.0000 |
| Head - StandingFoot | $-4.38$ | 3.54 | 1.0000 |
| Head - AverageShoes | $-7.50$ | 3.54 | 0.2040 |
| Hip - StandingFoot | $-6.25 \times 10^{-1}$ | 3.54 | 1.0000 |
| Hip - AverageShoes | $-3.75$ | 3.54 | 1.0000 |
| StandingFoot - AverageShoes | $-3.13$ | 3.54 | 1.0000 |
| Satisfaction | | | |
| Head - Hip | $-8.13$ | 5.00 | 0.6263 |
| Head - StandingFoot | $-6.88$ | 5.00 | 1.0000 |
| Head - AverageShoes | $-1.06 \times 10^{1}$ | 5.00 | 0.2022 |
| Hip - StandingFoot | 1.25 | 5.00 | 1.0000 |
| Hip - AverageShoes | $-2.50$ | 5.00 | 1.0000 |
| StandingFoot - AverageShoes | $-3.75$ | 5.00 | 1.0000 |
| Self-motion compellingness | | | |
| Head - Hip | $-1.25$ | 7.08 | 1.0000 |
| Head - StandingFoot | $-7.50$ | 7.08 | 1.0000 |
| Head - AverageShoes | $-1.13 \times 10^{1}$ | 7.08 | 0.6710 |
| Hip - StandingFoot | $-6.25$ | 7.08 | 1.0000 |
| Hip - AverageShoes | $-1.00 \times 10^{1}$ | 7.08 | 0.9453 |
| StandingFoot - AverageShoes | $-3.75$ | 7.08 | 1.0000 |
| VR Sickness | | | |

Continued on the next page ...

Continued from the previous page.

| contrast | estimate | SE | p.value |
|---|---|---|---|
| Head - Hip | $-2.21$ | 1.77 | 1.0000 |
| Head - StandingFoot | $-2.10$ | 1.77 | 1.0000 |
| Head - AverageShoes | $-4.29$ | 1.77 | 0.0923 |
| Hip - StandingFoot | $1.04 \times 10^{-1}$ | 1.77 | 1.0000 |
| Hip - AverageShoes | $-2.08$ | 1.77 | 1.0000 |
| StandingFoot - AverageShoes | $-2.19$ | 1.77 | 1.0000 |

End of the long table