

HeadJoystick: Improving Flying in VR using a Novel Leaning-Based Interface

Abraham M. Hashemian, Matin Lotfaliei, Ashu Adhikari, Ernst Kruijff, and Bernhard E. Riecke

Abstract—Flying in virtual reality (VR) using standard handheld controllers can be cumbersome and contribute to unwanted side effects such as motion sickness and disorientation. This paper investigates a novel hands-free flying interface—*HeadJoystick*, where the user moves their head similar to a joystick handle toward the target direction to control virtual translation velocity. The user sits on a regular office swivel chair and rotates it physically to control virtual rotation using 1:1 mapping. We evaluated short-term (Study 1) and extended usage effects through repeated usage (Study 2) of the HeadJoystick versus handheld interfaces in two within-subject studies, where participants flew through a sequence of increasingly difficult tunnels in the sky. Using the HeadJoystick instead of handheld interfaces improved both user experience and performance, in terms of accuracy, precision, ease of learning, ease of use, usability, long-term use, presence, immersion, sensation of self-motion, workload, and enjoyment in both studies. These findings demonstrate the benefits of using leaning-based interfaces for VR flying and potentially similar telepresence applications such as remote flight with quadcopter drones. From a theoretical perspective, we also show how leaning-based motion cueing interacts with full physical rotation to improve user experience and performance compared to the gamepad.

Index Terms—3D User Interface, Motion Sickness, Cybersickness, Flying, Travel Techniques, Virtual Reality

1 INTRODUCTION

FLYING has always been a fascinating dream for humanity, and despite current flying technologies such as planes, helicopters, paragliders, or wingsuits, flying is not yet easily accessible for most people. It also differs considerably from the long-held dream of bird-like, unencumbered and embodied flying experiences. As an alternative approach, virtual reality (VR) using head-mounted displays (HMDs) could provide a great opportunity to experience such embodied and unencumbered flying through virtual environments (VE), as VR can provide a first-person immersive and embodied experience. HMDs could also help provide a more compelling experience of flying in the real-world when used in telepresence/teleoperation scenarios, where the user controls an unmanned aerial vehicle (UAV), such as camera-equipped drones, and sees through its camera in real-time [1]. UAV telepresence can be used for different applications such as virtual aerial tourism [2], surveillance, inspection, or search and rescue in disaster areas [3].

Flying interfaces usually require the user to control different degrees of freedom (DoFs) for changing position (translation) and direction (rotation) of the simulated flying camera or actual UAV. For example, flying interfaces for

helicopters or quadcopters require controlling more DoFs (at least 4) than airplanes or fixed wing UAVs (at least 3), and thus allow for more control over the flight trajectory.

This paper investigates a simulated flying interface with four DoFs: forward/backward, up/downs, sideways, and yaw rotation, mimicking the controls used for quadcopter drones. Such an interface can be helpful in both simulations (e.g., video games and other VR applications) and telepresence applications (e.g., remote surveillance) due to its high maneuvering ability. For example, a well-designed 4DoF flying interface should allow users to reach their target position fast and accurately or rotate without translation to search for the next target position. A 4DoF flying interface could also help to control telepresence drones (which are predominately quadcopter-based) which allows the user to fly through pipes for inspection or through a wrecked building looking for survivors - chapter 8 of [4].

VR and telepresence flying applications share similar challenges when the user needs to control four DoFs, though. The standard flying interfaces for video games and VR (gamepad and hand-held controllers) and telepresence (i.e., proportional remote controls like radio-controlled aka RC controllers) essentially use two thumbsticks for locomotion control, and are usually cumbersome and require extensive training sessions for proficient control [5]. This motivated us to design a novel and more embodied and intuitive flying interface called “HeadJoystick”, aimed to reduce cognitive load compared to the standard handheld flying interfaces. HeadJoystick uses the head as a “joystick,” where users move their head (instead of deflecting the thumbstick) toward the target direction to control their simulated translation velocity. The user is seated on a regular office swivel chair and rotates it physically to control their simulated rotation using 1:1 mapping. This HeadJoystick was evaluated in two user studies focusing on short-term

- A.M. Hashemian, M. Lotfaliei, Ashu Adhikari, and B.E. Riecke are with the School of Interactive Arts & Technology, Simon Fraser University, Canada.
E-mail: {hashemia, mlotfali, ashua, ber1}@sfu.ca
- E. Kruijff is with the Institute of Visual Computing, Bonn-Rhein-Sieg University of Applied Sciences, Germany and the School of Interactive Arts & Technology, Simon Fraser University, Canada.
E-mail: ernst.kruijff@h-brs.de

Manuscript received XXX.XX, 2019; revised XXX.XX, XXXX; accepted XXX. XX, XXXX. Date of publication XXX. XX, XXXX.
For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below. Digital Object Identifier no. XX.XXXX/TVCG.XXXX.XXXXXXX

(Study 1) and extended usage effects (Study 2).

To this end, we designed a novel simulated drone racing task in HMD-based VR, where participants were asked to fly toward nine tunnel way-points and fly through the tunnels of decreasing diameter without colliding with the walls. Our test environment closely resembles operation of a UAV, to support transfer of our system and results to other usage domains besides standard VR environments. In our first study, 24 participants used four different interfaces to do this task, to tease apart the relative contributions of leaning-based translational cues versus full physical rotation cues: The *Gamepad*, which provided no physical motion cues beyond operating the thumbsticks, the *HeadJoystick* that provided leaning-based translational cues and full physical rotation cues, *RealRotation*, using the gamepad translation along with the chair physical rotation; and *LeaningTranslation*, using gamepad for rotation along with the leaning-based translation of the *HeadJoystick*. We measured performance, accuracy and precision and asked participants to compare these four interfaces in terms of different user experience aspects (e.g., enjoyment, presence, immersion, sensation of self-motion, preference) as well as usability measures (e.g., ease of learning, ease of use, motion sickness, task load). The second study was designed to investigate how results might generalize to extended exposure. To this end, a new set of 12 participants evaluated *HeadJoystick* versus *RealRotation* for doing eight rounds of the same 3D racing task. The main contributions of this study are:

- Introducing a novel low-cost leaning-based flying interface called *HeadJoystick*.
- Evaluating the *HeadJoystick* versus handheld controllers using a novel reach-the-target task combined with the tunnel-in-the-sky waypoint navigation task to comprehensively investigate diverse user experience, usability and the behavioral performance measures.
- Study 1 provides a deeper understanding of how leaning-based translation and full physical rotation each contribute to the overall user experience and performance.
- Study 2 investigates how repeated usage affects user experience and performance when using *HeadJoystick* versus handheld controllers, and corroborates the benefits of embodied (*HeadJoystick*) locomotion over hand-held controllers.

2 RELATED WORKS

In this section, we start with a general review of flying interfaces and then review flying interfaces similar to ours.

Various 4DoF flying interfaces have been investigated for immersive VR including hand-held interfaces [6], hand or arm-based gesture commands [7], [8], [9], [10], [11], [12], voice commands [6], [13], [14], and even brain-computer interfaces [15]. In general, these interfaces do not provide vestibular cues aligned with the visual motion direction of flight, which can reduce the believability of flying [16]. Moreover, the mismatch between visual and vestibular/proprrioceptive cues can cause or exacerbate visually induced motion sickness (VIMS), where the user feels motion sick without physically moving [17]. VIMS is known

as an unwanted side-effect in many virtual [18] or remote [19] flight systems, and will be referred to as simple motion sickness in the present work as it can also occur when users are physically moving.

We use the term embodied flying interfaces here to refer to interfaces that provide a visual 1st person perspective accompanied by at least some physical (including vestibular) self-motion cues. While HMDs can provide convincing visual cues of self-motion [20], it is not possible to provide full physical cues of self-motion without actual flying [16]. Therefore, embodied flying interfaces aim to create a believable flying experience by providing *limited* physical self-motion cues aligned with the vestibular/proprrioceptive sensory cues in an actual flight. These physical self-motion cues can be provided by the mechanical setups (such as in actuated moving-base flight simulators [21], [22]) or simply the user-powered body movements in leaning-based interfaces [5], [23], [24].

While several embodied flying interfaces use complex mechanical setups to provide physical self-motion cues to the user's body, we chose to design a leaning-based interface due to their simplicity and affordability for the majority of VR users. As an example of complex mechanical flying interfaces, moving-base flight simulators use motors/actuators to apply limited physical motion cues to the user's body [21]. Harnessing the user from ceiling is another fairly complex mechanical approach for embodied flying interfaces [25], [26], [27]. However, these mechanical interfaces usually have complicated setups and safety hazards, as summarized in [28]. *Birdly* is a mechanical interface for flying like a bird in VR [29] or telepresence applications [30], and applies limited physical motions to a user lying face-down on it. However, *Birdly* is too expensive (more than a hundred thousand dollars) for most VR home users, professionals, and UAV pilots.

2.1 Leaning-Based Interfaces

Leaning-based interfaces usually deploy user-powered leaning toward the target direction to control their simulated translation velocity without the need for any additional actuators. These interfaces generally use a velocity control paradigm, where the more the user leans, the faster they travel. While a seated user can lean their upper body and/or tilt the chair/stool they are sitting on [31], [32], [33], standing users can lean using their whole body [34], [35], [36]. In this section, we discuss leaning-based interfaces for 2D (ground-based) locomotion as they have been much more widely researched than 3D leaning-based interfaces, and also because our suggested interface (*HeadJoystick*) was originally designed for both 2D and 3D locomotion [37].

In this study, we investigate if leaning-based interfaces could be beneficial for flight (3D) control, given the diverse advantages of leaning-based over gamepad/joystick interfaces reported for ground-based (2D) locomotion. These advantages include an enhanced illusion of virtual self-motion (vection) [36], [38], [39], spatial perception and orientation [35], navigation performance [40], immersion and presence [34], [41], [42], enjoyment and engagement [34], [35], [36], as well as reduced motion sickness and cognitive load [40]. Additionally, leaning-based interfaces are hands-free,

which allow us to use our hands for other tasks (such as pointing, interacting with objects, or communicating) in VR and teleoperation applications, similar to how we can freely use our hands in the real world while walking [42], [43], [44], [45].

Leaning-based interfaces usually control the simulated rotations around the earth-vertical axis (yaw) either with the limited physical rotations using velocity control [31], [32], [33] or full physical rotations with 1:1 mapping between physical and simulated yaw rotations [34], [37], [40], [46]. Although limited rotation might be better for stationary displays such as projection screens, where the user cannot see the screen if they fully rotate, full physical rotation provides natural physical self-rotation cues and thus remove the visual-vestibular cue conflict for yaw rotations, which might lead to more believable self-motion experiences. However, they do require an HMD or 360 surround screens, or a screen rotating with the user as in moving-base motion simulators. Additionally, full physical rotation may help in reducing motion sickness compared to limited rotation due to reducing the conflict between visual and vestibular cues. Therefore, we use a full physical rotation approach for our interface, where the physical rotation of the user in the real world controls the direction of simulated camera using 1:1 mapping.

Allowing for full physical rotation can help users remain spatially oriented [47], [48], [49], [50], [51] by allowing them to more easily update their mental spatial orientation. Mixed results are reported about the importance of physical rotation for supporting spatial orientation when the user has no physical translation cues — as summarized in [52], [53]. However, some researchers reported that providing physical rotation with no or leaning-based translation could reach almost the same efficiency as actual walking in a navigational search task [40], [53]. While there can be challenges with too many rotations if a cabled HMD is used, this problem will soon lose relevance with the increasing quality and affordability of wireless HMDs or trackers entering the market. As an example, we used a wireless HTC-Vive HMD in our study.

2D Leaning-based interfaces have been designed for both standing users [34], [36], [54], [55], [56] and seated users [31], [32], [57]. For the current study, we chose a seated body posture due to comfort and safety reasons: As for comfort, seated users not only experience less discomfort, fatigue and leg-swelling in long-term usage [58], but they also experience less motion sickness compared to standing users [59] as predicted by postural instability theory [60]. Regarding safety, standing users might experience body sway during 3D virtual acceleration similar to VR roller coasters, and might fall and get hurt [61]. This motivated us to design a seated flying interface for the current study, even though our approach can easily be used for standing users as well if desired.

The aforementioned literature suggests that using a seated 4DoF flying interface with leaning-based translation and full physical rotation might be able to improve different aspects of 3D locomotion (e.g., vection, immersion, presence, enjoyment, and task-specific performance). However, there seems to be no prior published research that thoroughly investigate such an interface in terms of all these

aspects as far as the authors know, apart from studies that investigated partially similar interfaces in terms of limited aspects, as detailed in sections 2.2 and 2.3 below [5], [23], [24], [62]. Therefore, this gap in the literature motivated us to design HeadJoystick and evaluate it in terms of a wide range of aspects.

2.2 Leaning-Based Interfaces Controlling two DoFs

Table 1 compares the HeadJoystick with other leaning-based flying interfaces. In this section, we review leaning-based interfaces that control two DoFs, which are investigated for airplane control in virtual flight or fixed-wing drone control in remote flight. For example, Schulte *et al.* developed an upper-body leaning-based "dragon-riding" interface to control pitch and yaw of a simulated dragon [24] where a seated user leans backward or forward to pitch up or down respectively, and leans left/right to control their simulated yaw rotation. However, a dragon-riding interface might be unsuitable for most applications as the forward (translation) velocity was kept constant except when using a certain hand gesture to triple the speed for three seconds and then decelerating back to the normal speed. Dragon-riding interface was not compared with a standard controller such as RC remote controller or a gamepad.

Miehlsbradt *et al.* suggested a similar upper-body leaning-based interface - called "torso strategy", where the user moves their torso forward/backward and left/right to control the pitch and yaw/roll of a simulated fixed-wing airplane and thus fly up/down and turn left/right respectively [5]. In a virtual flight task, participants were asked to control a simulated fixed-wing drone and fly through a series of simulated waypoints. The results showed that torso-strategy outperformed standard RC remote controller and reached a performance level comparable to the Birdly flight simulator. Participants also used torso strategy to control a real quadcopter with constant forward velocity and no strafing, which reduced its DoFs similar to a fixed-wing drone. However, in that implementation users could not directly control translation velocity, and thus cannot really start or land or slow down, which makes it unfeasible for most realistic applications.

Rognon *et al.* also suggested a similar upper-body leaning-based interface to torso strategy — FlyJacket, where the user wears a backpack that supports their arms' weight and holds their arms up while the user was leaning [62]. The backpack was equipped with an inertial measurement unit (IMU), which enabled the user to lean forward/backward or left/right to control the pitch and yaw/roll of a drone, respectively. The participants were asked to fly a fixed-wing drone with constant forward velocity through several waypoints. Although FlyJacket had no significant improvement in performance compared to an RC remote controller, FlyJacket showed higher control on navigation, naturalness, and lower discomfort compared to the RC remote controller.

2.3 Leaning-Based Interfaces Controlling four DoFs

In this section, we review leaning-based interfaces that control four DoFs, which are investigated for VR applications or remote quadcopter control. Higuchi and Rekimoto [63] designed a telepresence interface called Flying Head, where

TABLE 1

Leaning-based flying interfaces. Note that all 2DoF interfaces used a fixed-wing (plane) locomotion paradigm, whereas the 4DoF interfaces used a quadcopter paradigm.

DoF	Body Posture	Interface	Rotation		Translation	
			Control	Rotation Input	Control	Translation Input
2	Seated	<i>Dragon-Riding</i> [24]	Velocity	Torso roll	Velocity	Torso pitch
2	Seated	<i>Torso-Strategy</i> [5]	Velocity	Torso yaw and position	Velocity	Torso position
2	Seated	<i>FlyJacket</i> [62]	Velocity	Torso roll	Velocity	Torso pitch
4	Standing	<i>Flying-head</i> [63]	Position	Head yaw	Position	HMD position
4	Standing	<i>Head-Rotation</i> [23]	Velocity	Head yaw	Velocity	Head pitch and roll
4	Standing	<i>Head-Translation</i> [23]	Velocity	Head yaw	Velocity	HMD position
4	Standing	<i>Modified Flying-head</i> [23]	Position	Torso yaw	Velocity	HMD position
4	Seated	<i>HeadJoystick</i>	Position	Chair yaw	Velocity	Head rotation center

a standing user controls the direction of the UAV with the direction of their head using 1:1 mapping, and the position of the UAV via the position of their head using 1:N mapping. Flying Head showed advantages over the joystick in two search and capture photo tasks in terms of ease of use, enjoyment, and the lower task completion time. However, because Flying Head uses a position control paradigm for simulated translation, the movement of UAV is limited to the user's head and body movements in the real world, which makes it not applicable to long-range flight and most realistic applications.

To the best of our knowledge, the only prior study that investigated leaning-based 4DoF flying interfaces and thus the most relevant prior work was done by Pittman and LaViola [23]: 18 participants flew through rectangular waypoints for about 90 seconds to compare a Wiimote interface similar to a gamepad with five other interfaces including three leaning-based flying interfaces: Head-Rotation, where the user controls drone translation by tilting their head forward and/or sideways; Head-Translation, where the user controls drone translation by moving their head forward/backward and/or sideways; and modified flying-head, where the user controls drone translation velocity by moving their head forward/backward and/or sideways, and controls drone rotation by rotating whole their body using 1:1 mapping. While results showed that the Wiimote interface performed best along almost all measures such as task completion time, comfort, ease of use, predictability, enjoyment, naturalness, and overall preference, the authors stated several technical issues that likely contributed to the general disfavor of leaning-based interfaces that motivated our studies: (1) *calibration*: 39% of participants reported low precision of leaning-based interfaces due to reasons such as incorrect calibration, thus we simplified the calibration process. (2) *Pose*: While all the interfaces were tested when users were standing, a number of users commented that using leaning-based interfaces could be easier when seated. As standing body posture could lead to higher discomfort, severe motion sickness, with more safety hazards compared to the seated body posture, we designed all our interfaces for seated users. (3) *zero-point*: Multiple participants mentioned drifting and difficulty to return to the zero point when using head-translation and modified flying-head, due to lack of visual feedback for the zero-point. Therefore, we asked our participants to set the zero-point when their back touches the chair backrest, so later they could easily find

this zero-point during flying without visual feedback. (4) *Technical issues*: Loss and oscillation of the drone's sensory information caused occasional stutter of the interface and side to side vibration of the drone during rotation when using modified flying-head. To address this, we used a virtual drone, which also allowed us to gradually reduce the size of waypoints (and thereby increased task difficulty) to study the achievable flying precision without and danger of crashing an actual drone.

While the aforementioned studies showed the potential of leaning-based interfaces for ground-based locomotion and 2DoF flying, it seems like leaning-based flying interfaces have not been investigated for 4DoF except the above-mentioned study [23], which had a few technical issues, and thus motivated us to design and conduct this study.

3 USER STUDIES

3.1 Research Questions

This study aims to thoroughly evaluate leaning-based 4DoF flying interfaces through 5 specific research questions:

RQ1: Do leaning-based interfaces improve user experience compared to hand-held controllers? 2D leaning-based interfaces are known to improve different aspects of locomotion experience including strongervection intensity [36], [38], [39], immersion and presence [34], [41], [42], as well as enjoyment [34], [35], [36]. As for leaning-based flying interfaces, while FlyJacket [62] improved user experience compared to hand-held interfaces, the head-rotation and head-translation interfaces in Pittman *et al.* were rated lower than hand-held devices in almost all aspects. However, since many studies reported improved user experience for ground-based leaning-based interfaces, we hypothesize that flying experience should also be improved by HeadJoystick.

RQ2: Do leaning-based interfaces improve flying performance compared to hand-held controllers? Embodied interfaces are known to improve locomotion performance compared to hand-held interfaces if they provide exact self-motion cues [64]. For example, bipedal walking for 2D locomotion or mimicking head movements in 3D locomotion (i.e., flying-head interface [63]) can improve locomotion performance. However, compared to hand-held interfaces, embodied interfaces that provide partial motion cues of locomotion have shown mixed results. Bowman *et al.*, reported reduced performance for partial motion cues [64]. Similarly, FlyJacket [62], flying-head, head-rotation,

and head-translation showed no significant improvements or lower performance compared to hand-held interfaces in a reach-the-target task [23]. Conversely, a torso-leaning-strategy showed higher performance than a hand-held device in recent studies of 3D flying controlling two DoF [5] and ground-based (2D) locomotion with 3 DoF control [40]. Given the technical issues of flying head, head rotation, and head translation to control a real drone [23], we hypothesize that the HeadJoystick should show similar results to the torso-leaning-strategy [5] and should improve performance compared to a hand-held controller.

RQ3: Can adding full physical rotation and leaning-based translation cues help to reduce visual-vestibular sensory conflicts and thus motion sickness? Providing full-translational sensory cues for flying is not possible unless the actual flying motions are replicated, as in isomorphic simulations [16]. Therefore, the maximum possible sensory data offered by an embodied flying interface (and without actually flying) could be full-rotational with partial-translational sensory data, similar to what the HeadJoystick offers. Considering that hand-held controllers provide minimal sensory data for both translation and rotation (in the form of haptic cues from the thumbsticks), evaluating our four interfaces allows us to investigate how minimal versus maximum-possible sensory data for the flight translation and rotation affects motion sickness.

The literature indicates mixed results in terms of how leaning-based interfaces affect motion sickness. For instance, some 2D locomotion studies reported that leaning-based interfaces did not reduce motion sickness compared to hand-held interfaces [34], [37], while others reported significant reductions of motion sickness using leaning-based interfaces [40]. Similarly, in 3D locomotion, flying-head, head-rotation, and head-translation did not reduce motion sickness using leaning-based interfaces [23], whereas FlyJacket reduced motion sickness [62].

As the sensory conflict theory of motion sickness [17], [18] suggests that reducing the cue conflict between different sensory cues indicating self-motion should reduce motion sickness, we predict that HeadJoystick (which was designed to reduce inter-sensory cue conflicts) should reduce motion sickness.

RQ4: How do leaning-based translation and full physical rotation each contribute to the overall user experience and performance? As far as the authors know, no prior research investigated how much leaning-based translation impacts the overall flying experience and/or performance with/without full physical rotation. Prior research on 2D (ground-based) navigation show mixed results regarding this research question (such as [52]). However, as full physical rotation could provide vestibular/proprioceptive sensory data similar to real-life like flying experience, we hypothesize that full physical rotation could improve the user experience and performance compared to limited/no physical rotation when using thumbsticks. As for the contribution of leaning-based translation without full rotation on the overall user experience and performance, there is mixed evidence: While Head-Translation [23] showed no improvement, FlyJacket [62] improved the user experience, and torso-strategy [5] improved performance. Due to the similarity with [5], [62], we predict that leaning-based trans-

lation in our study should improve both user experience and performance.

RQ5: How do user experience, usability, and performance change over repeated interface usage? Proficient control of handheld flying interfaces are known to require extended training sessions [5]. Prior research showed significant performance improvements during repeated usage of locomotion interfaces after a few trials in terms of speed [65], accuracy [66], number of errors [67], and the task completion time [34], [68]. Thus, we designed a second study to investigate how the findings of Study 1 which had relatively short exposure might or might not generalize to repeated and longer exposure. Especially, as motion sickness can build during continued exposure to VR - chapter 2.5 of [69], we aimed to investigate how motion sickness might change over extended usage of the leaning-based vs handheld interfaces. We hypothesized in RQ1-3 that using HeadJoystick improves user experience (RQ1) and performance (RQ2) and reduces motion sickness (RQ3) – here we hypothesize that these benefits of HeadJoystick will continue to hold even for extended usage. We addressed RQ1-4 primarily by Study 1, while Study 2 was designed to specifically address RQ5, and corroborate RQ 1-3 for repeated usage.

3.2 Task

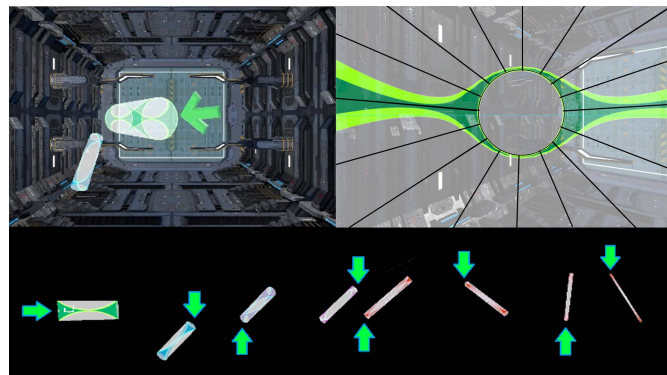


Fig. 1. Virtual environment used for the tunnel-in-the-sky task: flying through tunnels inside a spaceship hangar. Top: Environment from participant view, where the green arrow shows the entrance direction of the next tunnel. Middle: Environment from participant view, inside a tunnel. The black lines are added to show the cylindrical structure of the tunnel. Bottom: Side view of all tunnels showing how they get narrower. Green arrows show the entrance of each tunnel, illustrating the amount of required rotation to do this task.

A wide range of tasks have been used to evaluate flying interfaces, such as collecting objects [65], navigational search [70], pointing tasks [71], or capturing photos [63]. We chose *reach-the-target*, a well-known task in drone racing contests, where the user has to reach predetermined circular waypoints and fly through them [23], [24], [26], [27], [30], [71], [72]. Interface *accuracy* can be measured by the average distance from the desired path [73]. Since reach-the-target tasks have no predefined desired paths, we replaced the circular waypoints with a series of cylindrical *tunnels-in-the-sky* [74] that users were asked to fly through without colliding as illustrated in Figure 1. This allows us to quantify the interface accuracy as the average distance from the center of a tunnel when passing through it, because the most

optimal and safest way (i.e., least chance of collisions) to pass through a tunnel without collision should be the one where participants fly through its center in a fairly straight line.

As interface *precision* when navigating through tunnels depends on how much the interface allows the user to navigate through a narrow tunnel without collision [73], we also successively reduced the diameter of each tunnel, to make the task harder after passing each tunnel. The tunnel diameters were 6, 4, 3, 2.5, 2, 1.5, 1, and 0.5 meter (Figure 1, bottom). Participants were asked to fly through each tunnel in a specified direction without colliding with the tunnel walls. To impose precise flying, we penalized participants who collided with a tunnel's wall by asking them to fly through it again [24], which meant they had to fly around it to enter it again from the same side. This allowed us to use the average collisions per passed tunnels as a measure for the interface precision.

3.3 Virtual Environment

The virtual environment was designed as a flying practice inside a spaceship hangar as shown in Figure 1, to provide rich visual self-motion cues and a naturalistic visual reference frame. Tunnels were laid out such that users had to perform substantial rotations to get from one tunnel exit to the entrance of the next tunnel. Subsequent tunnels also differed in their yaw and pitch orientations to ensure that users needed to control their movement in different directions and had to control more than one DoF simultaneously to pass tunnels. To prevent participants from learning the path, the tunnels' layout was mirrored per trial horizontally and/or vertically in a randomized order. We also added green arrows to the entrance of the next activated tunnel to be sure that users knew where to go next. We also provided audio feedback to inform users if they passed or failed a tunnel.

3.4 Dependent Variables

To thoroughly evaluate our interfaces in a wide range of aspects, we selected a total of 15 dependent variables (DVs). They consisted of three behavioral performance measures, and 12 subjective DVs to measure six user experience factors and six usability aspects using an online questionnaire. As for behavioral measures [73], we recorded participants' performance during their flight in terms of *speed*, measured by the average time to pass a tunnel [65]; *accuracy*, measured by the average distance from the center of passed tunnels when flying through [5], [30], [74]; and *precision*, measured by the average number of collisions with the tunnel per passed tunnel.

We measured six user experience factors including the *SUS* questionnaire for spatial presence [75] with 6 questions on a Likert-based scale of 1-7; the first (and usually used) part of the *NASA-TLX* questionnaire with six questions to measure the task workload [76] on a continuous 0-100% scale.; and four questions with continuous answers between 0% to 100% including *enjoyment*, by asking how much participants enjoyed using each interface; *immersion*, by asking how much participants felt immersed i.e., captivated by the flying task; *vection intensity*, where 100% means that the

participant senses a compelling illusion of physical flight (self-motion) inside a stationary spaceship, while 0% means that the participant senses themselves stationary and the spaceship moves around them; and the *overall preference* by asking how much participants preferred the interface, where 0% means the worst interface, and 100% means the best interface they could imagine.

Our six usability measures consisted of the simulator sickness questionnaire (*SSQ*) [77] and five questions with a continuous answer between 0% to 100% including: *ease of use*, by asking how easy it was to use the interface; *ease of learning*, by asking how easy it was to learn using the interface; *long-term use*, by asking if the participant could imagine using the interface for a longer time than the study task; *daily use*, by asking to rate if they could imagine using the interface in daily applications; and the *overall usability*, by asking to rate the overall usability of the interface. A motion sickness (post-pre) score was defined by subtracting the total *SSQ* score obtained before exposure to any conditions from the total score obtained after exposure to each of the four conditions.

3.5 Apparatus

The virtual environments were presented using an HTC-Vive HMD with binocular field of view about 110° diagonally with a combined resolution of 2160 × 1200 pixels. The virtual environment was created using Unity3D 2018.2 and rendered on a dedicated PC (Intel Core-i7, Nvidia GTX-1060). The PC was connected to the HMD using a wireless TPCast adaptor to avoid entangling the HMD cable during physical rotations of participants (Figure 2). We attached the battery of the HMD wireless adaptor to the swivel chair and attached an additional Vive tracker to the chair backrest to measure chair orientation. We used a wireless Xbox-1 controller for the conditions that required a gamepad. Participants wore a noise-canceling headphone with an ambient sound of a spaceship to avoid distraction of possible background noises and to hear the audio cues if they passed or missed a tunnel.

3.6 Study 1

The goal of Study 1 was to investigate how using leaning versus thumbstick translation techniques, and physical versus thumbstick rotation techniques affects user performance and user experience (RQ1-4). Thus, we designed four different flying interfaces that differed in how a user controls translation and rotation. The techniques are named HeadJoystick, Gamepad, RealRotation, and LeaningTranslation, as shown in table 2. Each participant performed the task with all four interfaces. Due to our pilot tests, we limited the task completion time to 90 seconds (similar to the average task completion time in Pittman and LaViola [23]) to reduce the risk of severe motion sickness for inexperienced participants.

3.6.1 Locomotion Modes

This study compared four flying interfaces with different levels of physical motion cues for translation and rotation as illustrated in Figure 2 and Table 2, which are described below in more detail.

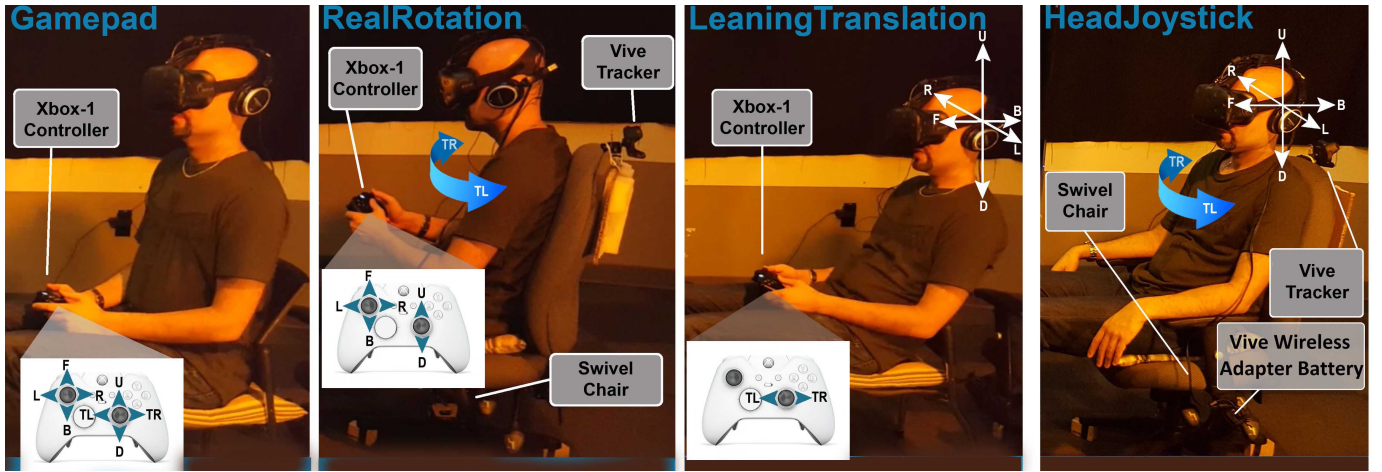


Fig. 2. All four flying interfaces compared in Study 1. Each interface controls flying along four degrees of freedom including forward(F)/backward(B), left(L)/right(R), up(U)/down(D), and turn-Left(TL)/turn-right(TR).

TABLE 2
Flying interfaces in Study 1, using color coding as in Figure 3

Rotation \ Translation	Gamepad-based	Chair-based
Gamepad-based	<i>Gamepad</i>	<i>RealRotation</i>
Head-based	<i>LeaningTranslation</i>	<i>HeadJoystick</i>

In the **Gamepad** condition, we used a classic controller scheme similar to [10], [23]. Participants moved the simulated camera forward/backward and sideways by pushing the left thumbstick forward/backward and sideways, respectively. The participants pushed the right thumbstick forward/backward and left/right to control Up/down movements and yaw rotations (left/right), respectively. The maximum translational velocity of the gamepad was 20m/s, the same as for all other interfaces. Based on pilot tests the maximum rotational velocity for the Gamepad and LeaningTranslation was set to $60^\circ/s$.

For **RealRotation**, participants translated the simulated camera using an Xbox-1 controller as in the gamepad condition, but rotated the simulated camera by physically rotating the office swivel chair they were seated on. We attached a Vive tracker to the backrest of the swivel chair to measure its yaw direction and mapped it to the yaw rotation of the simulated camera using a 1:1 mapping. For example, flying forward moved the simulated camera toward in the yaw direction of the swivel chair (not the head).

In the **LeaningTranslation** condition, participants rotated the simulated camera using the right thumbstick, but translated by moving their head toward the target direction. That is, the direction and distance of their head's position from its initial position (when starting flight) controls the direction and velocity of their simulated flight, which will be added to the position tracking. That is, for both LeaningTranslation and HeadJoystick conditions, we only consider the translation (not the rotation) of the users' head to control the simulated translation. As none of our interfaces consider the direction of the user's head to control the simulated rotation or translation, users could rotate their head freely to see

the virtual environment without affecting their simulated self-motion. The motion control model details are discussed in the appendix.

For **HeadJoystick**¹, simulated rotation was controlled by the physical rotation of the chair as in the RealRotation condition. Participants controlled the simulation translation using head movements similar to the LeaningTranslation interface with one difference: While LeaningTranslation uses a static zero-point (initial position of the head), HeadJoystick uses a dynamic zero-point to compensate for chair movements. That is, HeadJoystick uses the position and orientation of the chair-attached Vive tracker to continually update the position and orientation of the zero point, to keep it stationary with respect to the chair (not the room). In other words, the user could always find the zero point and stop the simulated translation easily by sitting upright and touching the chair backrest, even after rotating the chair or accidentally moving it on the floor. Dynamic zero point allows the user to rotate without translating even if the global position of their head changes during the yaw rotation of the chair. The HeadJoystick motion details are discussed in the appendix.

3.6.2 Participants

We recruited 24 students (12 females) between 19-50 years old ($M = 25.6$, $SD = 6.3$) for this study. 33% of participants had no prior experiences with HMDs, and 50% of them reported that they play video games on a daily or weekly basis using either online 3D PC games or gaming consoles. None of them had previous experience with any of our interfaces except the gamepad, which all of them were familiar with. Two additional participants did not finish the study due to motion sickness and were thus excluded from data analysis. We compensated participation time by either course credit or 15 CAD\$ for a 75 minutes experiment. The local ethics board approved this research (#2015s0283).

3.6.3 Experimental Design

This within-subject study compared gamepad control of a virtual drone with three more embodied interfaces that used

1. Video for HeadJoystick (<https://youtu.be/zVODu2ARV54>)

TABLE 3

Analysis of variance results for all dependent variables of the Study 1: Significant effects ($p \leq 5\%$) are written in bold, and were always in the direction of enhanced user experiences for embodied versus gamepad translation/rotation. The effect strengths partial Eta squared (η_p^2) indicates the percentage of variance explained by a given factor.

	Embodied Translation (yes/no)			Embodied Rotation (yes/no)			Interaction (Translation-Rotation)		
	F(1,23)	p	η_p^2	F(1,23)	p	η_p^2	F(1,23)	p	η_p^2
Enjoyment	50.8	<0.001	0.688	8.50	0.008	0.27	3.61	0.07	0.136
Preference	45.4	<0.001	0.664	14.3	0.001	0.383	2.08	0.16	0.083
Immersion	26.8	<0.001	0.538	7.66	0.011	0.25	0.056	0.815	0.002
Vection Intensity	13.7	0.001	0.373	4.29	0.05	0.157	0.098	0.757	0.004
Long-Term Use	12.6	0.001	0.353	9.18	0.006	0.285	0.761	0.392	0.032
Daily Use	16.5	<0.001	0.418	7.41	0.012	0.244	1.53	0.229	0.062
Overall Usability	27.9	<0.001	0.549	6.30	0.02	0.215	0.907	0.351	0.038
Presence (SUS)	20.1	<0.001	0.466	3.21	0.087	0.122	0.756	0.394	0.032
Ease of Use	16.6	<0.001	0.42	0.035	0.853	0.002	9.67	0.005	0.296
Ease of Learning	11	0.003	0.324	0.013	0.908	0.001	5.53	0.028	0.194
NASA-TLX	16.6	<0.001	0.419	1.57	0.223	0.064	7.99	0.01	0.258
Absolute Distance Error	70.4	<0.001	0.754	0.015	0.904	0.001	0.462	0.503	0.02
	<i>Z</i>	<i>p</i>		<i>Z</i>	<i>p</i>				
Motion Sickness (post-pre)	1.47	0.141		2.49	0.013				
Passed Tunnels	4.30	<0.001		1.55	0.120				
Collisions	3.89	<0.001		1.90	0.057				

either leaning-based translation (“LeaningTranslation”), full physical rotation (“RealRotation”), or both (“HeadJoy-stick”). Each participant completed 4 practice trials and 4 main trials, consisting of a factorial combination of 2 translation modes {embodied, gamepad} \times 2 rotation modes {embodied, gamepad}. Each main trial was preceded by a practice trial and only data from the main trial was analyzed, as the length of practice trials varied per participant, and we wanted to compensate for initial learning effects. Interface conditions were counterbalanced across participants using a Latin-square design.

3.6.4 Procedure

After reading and signing the informed consent form, participants filled an initial SSQ questionnaire of motion sickness [77]. Then each participant performed the fly-through-tunnels-in-the-sky task for each of the four interface conditions. Participants completed two trials per interface: a *practice trial*, where participants practiced the interface and flew through as many tunnels as they could until they felt comfortable, or one minute passed, whichever came first; This was immediately followed by a *main trial*, where participants had 90 seconds to fly through as many tunnels as they could. After completing the main trial with each interface, participants were asked to answer two Likert-based questionnaires including SSQ and other usability and user experience measures to evaluate the interface. Answering these questionnaires also provided participants a resting time before they used the next interface. After finishing all four interfaces, we explored reasons behind participant’s answers in a semi-structured interview.

3.6.5 Results

Data were analyzed using 2×2 repeated-measures ANOVAs for the independent variables embodied translation {yes/no} and embodied rotation {yes/no}, and

Tukey post-hoc tests for pairwise comparisons. We applied Greenhouse-Geisser correction when the sphericity assumption was violated. We analyzed ordinal data (i.e., number of passed tunnels) and ratio data that violated the normality assumption in Shapiro-Wilkes test (i.e., average collisions per passed tunnels and motion sickness post-pre scores) using Wilcoxon signed-rank test for main effects of embodied translation and embodied rotation. Due to the large number of DVs, we summarized main effects and interactions in Table 3, with post-hoc results presented together with descriptive statistics in Figure 3.

Main effects and interactions: Providing **embodied (head-based) translation** showed a significant main effect and positively affected 14 measures (all but motion sickness) compared to the gamepad translation (see Table 3). As for the user experience factors, embodied translation yielded significantly increased enjoyment, higher spatial presence (SUS questionnaire mean), improved immersion, stronger vection intensity, higher preference ratings, and reduced task load (NASA-TLX scores). As for the usability measures, embodied translation also yielded significant benefits in terms of being easier to use, easier to learn, longer-term use, more potential for daily usage, and enhanced overall usability. As for the performance measures, embodied translation yielded significantly increased accuracy (decreased absolute distance error), as well as in increased number of passed tunnels, and reduced collisions.

Providing **embodied (physical) rotation** also showed significant main effects and improvements compared to gamepad rotation in eight out of 15 DVs (see Table 3). As for the user experience factors, embodied rotation yielded significantly increased enjoyment, improved immersion, enhanced vection intensity, and higher overall performance ratings. As for usability measures, embodied rotation also yielded significantly enhanced overall usability, longer-term use, and more potential for daily usage, while also reducing

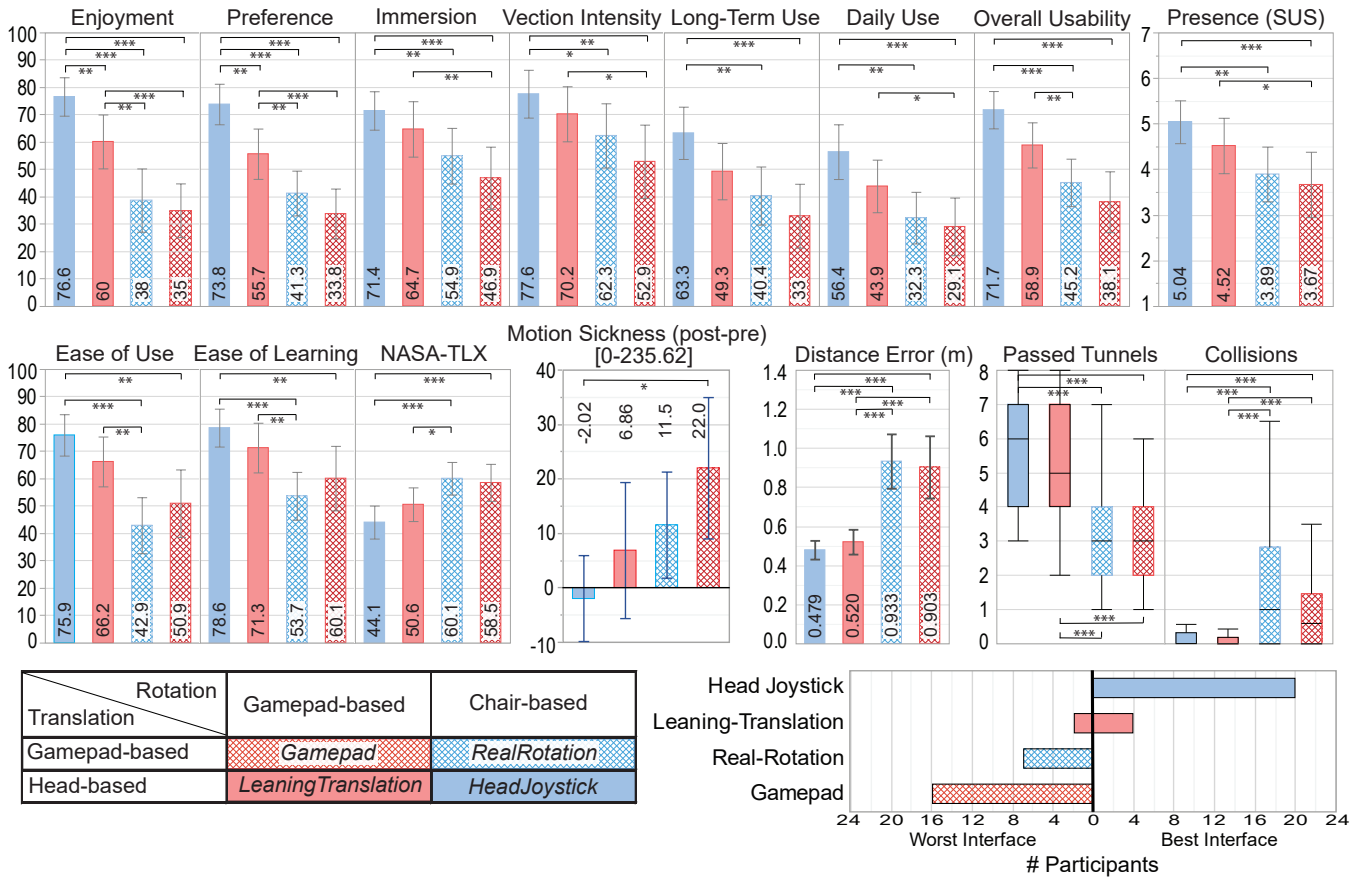


Fig. 3. Study 1 results: Mean data of user experience and performance measures except *Passed Tunnels* and *Collisions* plots, which show medians. Error bars indicate confidence intervals ($CI = 95\%$), annotated bars represent significance levels of post-hoc and non-parametric tests (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$). The bottom right plot shows the participants' rankings of best and worst interface from the post-experiment interview. For one participant, who chose both Gamepad and RealRotation as their worst interface, we included their answer in both ratings.

motion sickness. However, embodied rotation did not show a significant effect compared to the gamepad in terms of accuracy (absolute distance error), ease of use, ease of learning, task load, passed tunnels, and collisions. As for the absolute motion sickness levels, highest total SSQ scores were reported after using Gamepad ($M = 48.6, SD = 41.9$), followed by RealRotation ($M = 43.2, SD = 38.0$), then LeaningTranslation ($M = 41.1, SD = 36.3$), and finally HeadJoystick ($M = 31.5, SD = 27.7$). Note that for the ANOVA and Figure 3 we used the difference between post and pre-scores instead of the absolute SSQ values to avoid carryover effects.

An **interaction between translation and rotation** qualified these main effects for three (out of 12) DVs including ease of use, ease of learning, and task load (NASA-TLX) as illustrated in Table 3 and Figure 3. That is, the effect of adding embodied (leaning-based) translations depended on whether rotations were performed by gamepad or physical rotations: when rotations were controlled by gamepad (red bars in Figure 3), switching to leaning-based embodied translations instead of gamepad translations provided no significant benefit for these measures (see also post-hoc analysis in Figure 3). Conversely, when virtual rotations were controlled by physical rotations (blue bars in Figure 3), switching to leaning-based embodied translations instead of gamepad translations provided more substantial and

significant benefits in terms of increased ease of use and ease of learning, and reduced task load. To investigate if prior gaming experience improves performance, we conducted an additional ANOVA with the added between-subject factor of **prior gaming experience** {yes, no}. Participants who played 3D first-person games on a daily or weekly basis passed more tunnels ($M = 5.10\%, SD = 1.94\%$) compared to non-gamer participants ($M = 3.35\%, SD = 1.79\%$), $F(1, 19.5) = 13.5, p = 0.002, \eta_p^2 = 0.381$. The performance of the participants with prior gaming experience was consistently better with every interface. Participants who played 3D first-person games on a daily or weekly basis also rated interfaces easier to learn ($M = 71.5\%, SD = 23.9\%$) compared to non-gamer participants ($M = 60.3\%, SD = 22.4\%$), $F(1, 22) = 4.87, p = 0.038, \eta_p^2 = .181$. The participants with prior gaming experience consistently rated all the interfaces easier to learn compared to non-gamer participants. Prior gaming experience showed no significant interactions or effects on any other DVs.

Post-hoc pairwise comparisons: **HeadJoystick** showed significant benefits in pairwise comparisons compared to both the RealRotation and Gamepad conditions in most of our 15 DVs (see Figure 3). The only exception was motion sickness, where using the HeadJoystick reduced motion sickness only compared to the gamepad, but not the RealRotation condition. That is, compared to RealRotation and Gamepad

conditions, the HeadJoystick significantly increased enjoyment, preference, immersion, vection intensity, long-term use, daily use, overall usability, spatial presence, ease of use, ease of learning, and the number of passed tunnels, while reducing task load, absolute distance error, and average number of collisions. Compared to LeaningTranslation, HeadJoystick showed significantly higher enjoyment and preference. The other dependent measures showed only trends in the same direction that did not reach significance. In the post-experiment interview, 20 out of 24 participants (83%) chose HeadJoystick as the best (most favorite) interface, as illustrated in the bottom right plot of Figure 3.

LeaningTranslation showed significant benefits compared to using RealRotation and Gamepad in terms of nine out of 15 dependent measures (see Figure 3). Compared to using the RealRotation, LeaningTranslation yielded significantly increased number of passed tunnels, enjoyment, preference, overall usability, ease of use, ease of learning, with a reduced task load, absolute distance error, and average number of collisions. Compared to using the Gamepad, LeaningTranslation showed significantly increased number of passed tunnels, enjoyment, preference, immersion, vection intensity, daily use, spatial presence, as well as decreased absolute distance error and average number of collisions. In the post-experiment interview, 4 out of 24 participants (17%) chose the LeaningTranslation as the best (most favorite) interface while 2 participants (8%) chose it as the worst (least favorite) interface.

RealRotation did not show significant differences compared to the Gamepad in any of the 15 dependent measures, indicating that providing real rotations alone does not provide any benefits when translations are still controlled by gamepad (instead of leaning). In the post-experiment interview, 16 participants (67%) chose the gamepad and seven participants (29%) chose RealRotation as worst (least favorite) interface, while no participant chose any of them as the best (most favorite) interface.

3.6.6 Discussion

Study 1 provided evidence for the advantages of leaning-based over gamepad translation in terms of all user experience factors, usability aspects, and performance measures. However, using each interface only for 90 seconds might not be enough for a thorough evaluation of the interfaces, especially given that handheld flying interfaces (such as gamepad or RC controller) are known to require longer periods of time to be used efficiently [5]. Moreover, due to the short duration of Study 1, participants' subjective responses might have been influenced by the novelty aspect of the embodied interfaces, which might change for prolonged or repeated usage. Study 2 was designed to address these concerns and gain a deeper understanding of how user experience, usability, and performance might change during repeated exposure, and if the observed benefits of the leaning-based interface (HeadJoystick) might replicated and generalize to extended usage without increasing motion sickness critically.

3.7 Study 2

Study 2 was designed to address RQ5 and investigate how usability, user experience, and performance might change

over repeated interface usage, and if the observed benefits of leaning-based interfaces such as the HeadJoystick would generalize to multiple repetitions of the task. Repeated interface usage was expected to address initial learning effects and increase familiarity, which might benefit both the dual-thumbstick control scheme and the HeadJoystick which was a new interface for all participants. The overall experimental design and procedure of Study 2 was the same as for Study 1 apart from the changes described below.

Comparing leaning- vs. thumbstick translation: To reduce the potential for motion sickness, we excluded the two conditions from the first study that used thumbstick rotation, and only compared the two conditions using full physical rotation, where translations were controlled either by leaning (HeadJoystick) or thumbstick (RealRotation).

Eight trials per interface: Instead of one 90s trial per interface, we asked each participant to fly eight trials of 60s per interface to investigate how the different measures change over time due to learning/exposure effects. As our pilot studies showed some participants getting motion sick and dropping the experiment before completion, trial duration was reduced to 60s to reduce overall experiment duration while allowing for detection of learning/exposure effects.

Post-trial questionnaire: After each trial, we asked participants to verbally rate their motion sickness as well as perceived task difficulty on a 0-100% scale.

Reduced maximum velocity: As users in pilot studies stated that the controller thumbsticks were too sensitive and might induce severe motion sickness after a few trials, we reduced the maximum speed from 20 to 8 m/s, to reduce motion sickness and increase the usability of the thumbsticks.

Smooth acceleration: Based on user feedback about increased motion sickness during abrupt speed changes, we limited the possible accelerations/decelerations using Unity's SmoothStep function (see appendix), resulting in smoother velocity profiles (almost like inertia). Limiting accelerations was intended to reduce visual-vestibular cue conflict and the potential for motion sickness, and make the flying experience more realistic.

Using controller instead of Gamepad: As most VR HMDs deploy two separate controllers for each hand instead of a gamepad, we asked participants to use two Valve Index controllers, which have a similar-sized thumbstick as the gamepad used in Study 1. To avoid confusion, we call this RealRotation condition in Study 2 the "Controller" condition. We used a thumbstick mapping similar to the gamepad in the RealRotation condition of Study 1, where the left thumbstick controls forward/backward and sideways and the right thumbstick controls elevation.

Similar velocity transfer function for both conditions: To address the feedback from Study 1 participants that lower speeds were harder to control with the thumbsticks (which used linear mappings in Study 1), we used the same exponential transfer function for both thumbs-stick and leaning-based velocity control in Study 2.

3.7.1 Participants

We recruited 12 graduate students (5 females) between 25-37 years old ($M = 30.1, SD = 3.53$) for this study. Six participants (50%) had no prior experiences with VR HMDs, six of them (50%) reported playing 3D (first-person view)

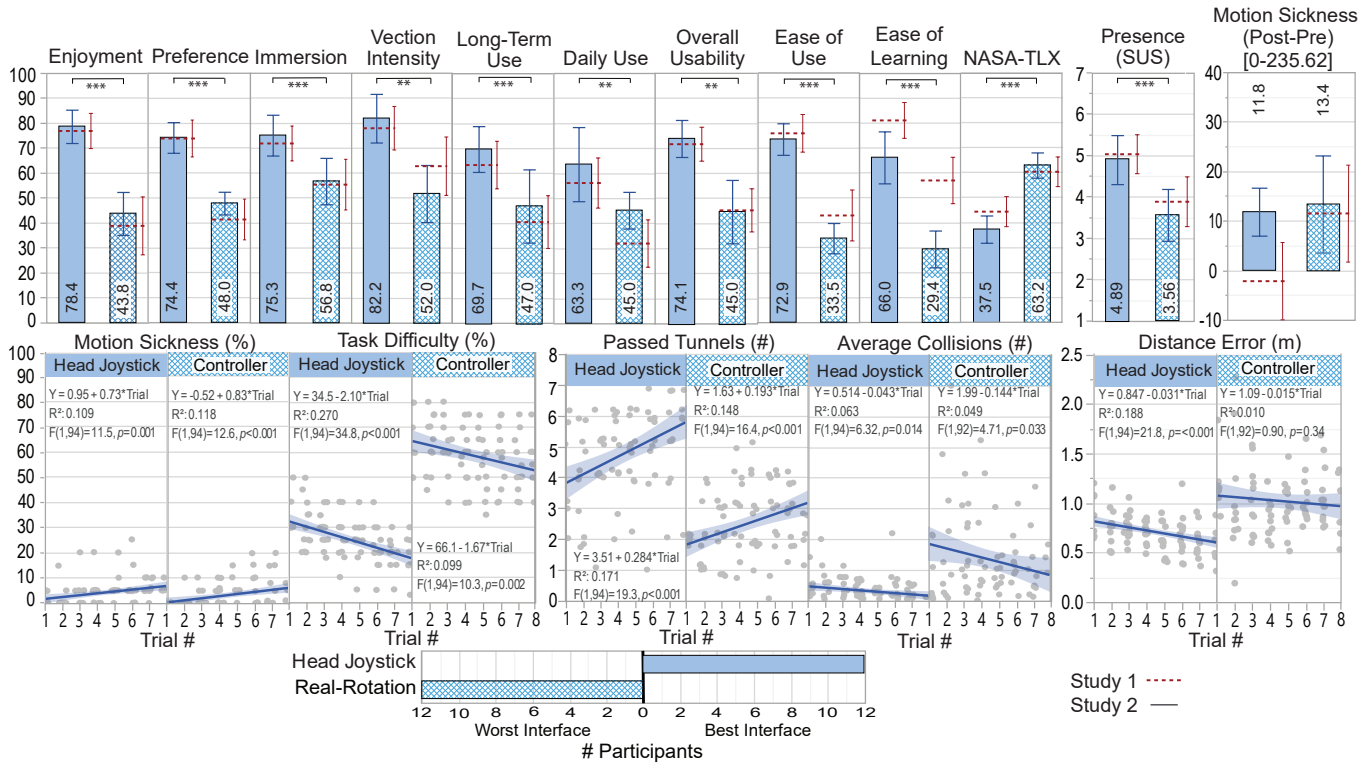


Fig. 4. Study 2 results of comparing HeadJoystick (in solid blue) versus Controller (in hatched blue): Top row shows mean data of user experience, usability, and performance measures. Error bars indicate confidence intervals ($CI = 95\%$), annotated bars represent significance t-tests differences ($* p < 0.05$, $** p < 0.01$, $*** p < 0.001$). Means from Study 1 are added as red dashed lines for easier comparability. Middle row shows user experience and performance changes over trials, including linear regression results. Gray dots indicate individual participants data and are jittered to improve visibility. The bottom plot shows the participants’ rankings of best and worst interface from the post-experiment interview.

video games on a daily or weekly basis. They had no prior experience with our interfaces, and we compensated their time for a 75 minutes experiment by offering a chance to try VR games for a couple of hours. The local ethics board approved this research (#2015s0283).

3.7.2 Results

TABLE 4

T-test results for dependent variables of Study 2: Significant effects ($p \leq 5\%$) are written in bold, and were always in the direction of enhanced user experiences for HeadJoystick over Controller. The effect size *Cohen’s d* indicates the magnitude of effect i.e., the difference between two means expressed in standard deviations.

	t(23)	p	Cohen’s d
Enjoyment	15.0	<0.001	4.33
Preference	11.5	<0.001	3.32
Immersion	4.89	<0.001	1.42
Vection Intensity	4.83	0.001	1.40
Long-Term Use	5.39	<0.001	1.55
Daily Use	3.21	0.008	0.924
Overall Usability	3.98	0.002	1.15
Presence (SUS)	6.81	<0.001	1.96
Ease of Use	12.4	<0.001	3.58
Ease of Learning	7.21	<0.001	2.08
NASA-TLX	7.63	<0.001	2.20
Motion Sickness (Post-Pre)	0.114	0.742	0.099

We compared HeadJoystick with Controller by analyzing 12 DVs using two-tailed repeated measures (paired) t-

tests as the data did not violate the Shapiro-Wilk test of normality. Due to the large number of DVs, we summarized t-test results in Table 4 and descriptive statistics in Figure 3. HeadJoystick showed significant benefits over Controller in 11 of our 12 measures (see top row in Figure 4) except motion sickness, which showed no significant difference. Total motion sickness scores were overall relatively low after using both the Controller ($M = 27.7, SD = 19.4$) and the HeadJoystick ($M = 25.6, SD = 18.9$). That is, compared to Controller, HeadJoystick yielded significantly increased enjoyment, preference, immersion, vection intensity, long-term use, daily use, overall usability, spatial presence, ease of use, and ease of learning, while reducing task load. Effect sizes (Cohen’s d) were large ($d \geq 0.8$) for all significant effects, indicating substantial benefits of the HeadJoystick even for prolonged usage, corroborating findings from Study 1.

To investigate if **prior gaming experience** improves performance, we conducted an additional ANOVA with the added between-subject factor of prior gaming experience {yes, no} and interface as the within-subject factor. Participants who played 3D first-person games on a daily or weekly basis passed more tunnels ($M = 4.29\%, SD = 1.69\%$) compared to non-gamers ($M = 3.19\%, SD = 1.74\%$), $F(1, 10) = 5.42, p = 0.042, \eta_p^2 = 0.331$. None of the other DV showed any significant effects of gaming experience, though, and there were no significant interactions.

To analyze how HeadJoystick and Controller affect the per-trial measures of motion sickness, task difficulty, number of passed tunnels, and average collisions over time,

we used 2×8 repeated-measures ANCOVAs for the independent variables interface and trial number. We analyzed ordinal trial data (i.e., number of passed tunnels) and ratio data that violated the normality assumption in Shapiro-Wilk tests (i.e., average collisions per passed tunnels and motion sickness) as rank-transformed data. We applied Greenhouse-Geisser correction when the sphericity assumption was violated. We summarized correlation results in the middle row of Figure 4. First and last trials were compared using planned contrasts.

Motion sickness showed a significant main effect of trial number ($F(1, 11) = 11.9, p = 0.005, \eta_p^2 = 0.459$), and linear regressions in Figure 4 corroborated significant increases in motion sickness over time for both the HeadJoystick ($p = 0.001, R^2 = 0.109$) and Controller ($p < 0.001, R^2 = 0.118$). Interface did not show any significant main effect or interaction with trial number. Motion sickness was overall low ($M = 3.73\%, SD = 5.35\%$) and increased from the first to the last trial from 2.08% ($SD = 3.34\%$) to 5.92% ($SD = 2.87\%$) for the HeadJoystick ($p = 0.041$), and from 0.42% ($SD = 1.44\%$) to 5.75% ($SD = 8.36\%$) for the Controller ($p = 0.007$).

Task difficulty was rated as overall lower for the HeadJoystick ($M = 25.0\%, SD = 9.31\%$) than the Controller ($M = 58.6\%, SD = 12.3\%$), $F(1, 11) = 114, p < 0.001, \eta_p^2 = 0.598$, and showed a significant main effect of trial number $F(1, 11) = 47.8, p < 0.001, \eta_p^2 = 0.229$. Linear regressions in Figure 4 indicate that task difficulty ratings decreased significantly over the course of the eight trials for both HeadJoystick ($p < 0.001, R^2 = 0.270$) and Controller ($p = 0.002, R^2 = 0.099$). More specifically, task difficulty decreased between the first and last trial from 31.3% ($SD = 8.56\%$) to 17.8% ($SD = 7.25\%$) for the HeadJoystick, ($p < 0.001$), and from 65.8% ($SD = 12.4\%$) to 53.7% ($SD = 11.5\%$) for the Controller ($p < 0.001$). There was no significant interaction between interface and trial number.

Performance was assessed in terms of the number of tunnels participants managed to pass in each trial, the number of collisions with the tunnel walls per passed tunnels, and the average distance error from the tunnel center while passing through. The **number of passed tunnels** showed a significant main effect of interface ($F(1, 11) = 53.4, p < 0.001, \eta_p^2 = 0.280$), with overall more tunnels passed for the HeadJoystick ($M = 4.79, SD = 1.57$) compared to the Controller ($M = 2.50, SD = 1.15$). Trial number showed also a significant main effect ($F(1, 11) = 62.2, p < 0.001, \eta_p^2 = 0.188$), which was qualified by a significant interface-trial interaction ($F(1, 11) = 6.47, p = 0.027, \eta_p^2 = 0.028$). As illustrated in Figure 4 and the linear regressions, this indicates significant performance improvement over the trials for both the HeadJoystick ($p < 0.001, R^2 = 0.171$) and the Controller ($p < 0.001, R^2 = 0.148$). Between the first and last trials the number of passed tunnels increased from 3.75 ($SD = 1.36$) to 5.83 ($SD = 1.53$) for the HeadJoystick ($p < 0.001$), and from 1.75 ($SD = 0.622$) to 3.08 ($SD = 1.08$) for the Controller ($p < 0.001$). The significant interaction suggests that the performance improvement was larger for the HeadJoystick compared to the Controller, which is corroborated by the steeper slope of the linear regression fit in Figure 4.

The **number of collisions per passed tunnel** showed a similar performance benefit (reduced collisions) for the HeadJoystick ($M = 0.319, SD = 0.398$) compared to the Controller ($M = 1.34, SD = 1.50$), $F(1, 11) = 15.6, p = 0.002, \eta_p^2 = 0.167$. There was also a significant main effect of trial number $F(1, 11.1) = 5.13, p = 0.045, \eta_p^2 = 0.071$, with collisions decreasing between the first and last trials from 0.413 ($SD = 0.500$) to 0.093 ($SD = 0.160$) for the HeadJoystick ($p = 0.505$), and from 1.83 ($SD = 2.40$) to 1.08 ($SD = 0.704$) for the Controller ($p = 0.125$). This performance improvement over the course of the eight trials was corroborated by significant negative linear regressions for both the HeadJoystick ($p = 0.014, R^2 = 0.063$) and the Controller ($p = 0.033, R^2 = 0.049$). There was no significant interaction.

The **average distance error** also showed a similar performance benefit (reduced distance error) for the HeadJoystick ($M = 0.709m, SD = 0.163m$) compared to the Controller ($M = 1.02m, SD = 0.353m$), $F(1, 10.7) = 30.4, p < 0.001, \eta_p^2 = 0.080$. There was also a significant main effect of trial number $F(1, 10.2) = 18.6, p = 0.002, \eta_p^2 = 0.076$, with distance error decreasing between the first and last trials from 0.839m ($SD = 0.131m$) to 0.615m ($SD = 0.125m$) for the HeadJoystick, ($p = 0.013$), and from 1.14m ($SD = 0.322m$) to .935m ($SD = 0.255m$) for the Controller ($p = 0.020$). This performance improvement over the course of the eight trials was corroborated by significant negative linear regressions for the HeadJoystick ($p < 0.001, R^2 = 0.188$), whereas the Controller showed no significant linear decrease in distance error ($p = 0.344, R^2 = 0.010$). There was no significant interaction.

The top row of Figure 4 illustrates that the observed differences between HeadJoystick and RealRotation/Controller showed similar data patterns (benefits for HeadJoystick) for both short-term usage (90s in Study 1, indicated as red dashed lines) and extended (repeated) usage in Study 2 (8 trials). Even the actual values were relatively similar between Study 1 and 2 for almost all subjective measures including enjoyment, preference, immersion, vection intensity, long-term use, daily use, overall usability, ease of use, task load, and presence. This is confirmed by running exploratory 2×2 ANOVAs with the factors Study {1 vs. 2} and interface {HeadJoystick vs. RealRotation/Controller}, which showed no significant main effects of study for any of these DV. Only ease of learning showed overall lower ratings in Study 2 vs. 1 ($p < 0.001$). There were, however, significant interactions between study and interface for vection intensity ($p = 0.044$), ease of learning ($p = 0.001$), and task load ($p = 0.024$), indicating more pronounced differences between HeadJoystick and RealRotation/Controller for extended usage in Study 2 vs. 1. Performance measures all showed improvements over repeated trials (Figure 4 middle-row), suggesting learning/practice effects as expected. In fact, after 8 trials participants in Study 2 managed to pass about as many tunnels in a 60s trial as participants in Study 1 in a 90s trial.

In the post-experiment interview, all 12 participants chose HeadJoystick as the best (most favorite) interface, which is shown in the bottom plot of Figure 4.

4 GENERAL DISCUSSION

Both studies showed conclusive evidence for the advantages of leaning-based over thumbstick translation in general, and specifically HeadJoystick over handheld controllers in terms of most of the user experience factors, usability aspects, and performance measures. In the remainder of this section, first we discuss results of Study 1 in the context of our research questions RQ1-RQ4 and discuss potential reasons for the observed effects. Then we discuss short-term vs repeated usage effects of our interfaces in the context of RQ5 based on data from Study 2. Therefore, unless stated otherwise, we refer to Study 1 results when discussing RQ1-4, and refer to Study 2 results when discussing RQ-5.

4.1 RQ1: Leaning-based interfaces improve user experience

Results confirmed our hypothesis that leaning-based interfaces improve different aspects of user experience compared to using thumbsticks (see Figure 3). While previous research showed improved naturalness and control over flight trajectory when using a 2DoF leaning-based interface such as FlyJacket [62], our findings extend knowledge by providing more thorough and conclusive evidence that 4DoF leaning-based interfaces can indeed improve a wide range of measures related to the user experience both in short-term (Study 1) and repeated usage of the interface (Study 2). Note that results patterns have been fairly consistent across Study 1 and 2, and effect sizes of all significant effects were all large ($\eta_p^2 > 0.14$ and Cohen's $d > 0.8$), and p -values were relatively small ($p < 0.008$), suggesting that effects (and the benefits of leaning-based interfaces) are substantial and not likely to be caused by false positives due to testing multiple measures. If anything, repeated usage of the interface in Study 2 showed more pronounced advantages of the HeadJoystick over Controller, indicating that the benefits observed in Study 1 generalize to more extended usage, and were not caused by initial novelty or first-exposure effects.

Compared to prior works, our conclusive results suggests that previously reported disadvantages of 4DoF (flying) leaning-based interfaces such as head-rotation and head-translation [23] might have originated from technical issues as discussed in section 2.3. For 2D (ground-based) locomotion, prior research showed benefits of leaning-based interfaces over hand-held controllers in terms of increased vection intensity [36], [38], [39], higher immersion and presence [34], [41], [42], and increased enjoyment [34], [35], [36]. Our findings show that these advantages can, in fact, generalize to 3D (flying) locomotion. Moreover, our results show additional advantages of leaning-based 3D interfaces in terms of usability measures such as ease of use, ease of learning, task load, long-term use, and daily use.

In the post-experiment interview, eight participants mentioned that HeadJoystick allowed for the most realistic experience of being in and moving through the virtual environment. For example, participants stated *"It [HeadJoystick] felt real. I am afraid of height, and using HeadJoystick, I could actually feel the height"*(p13), *"When I have more body motion, it feels like I am in a space station, but gamepad feels more like I am in a game"*(P8). The improved user experience and usability of HeadJoystick over thumbsticks may be

due to the alignment of head translation direction (and associated vestibular and proprioceptive cues) with the resulting simulated translation. In fact, HeadJoystick was designed to mimic real-world self-motion cues during the movement initiation (initial acceleration), where we lean a bit in the direction of intended travel before taking a step in that direction. Note that most previous leaning-based seated interfaces used weight-shifting (e.g., dragon-riding [24]), upper-body deflection (e.g., torso-strategy [5]), and/or tilting the chair/stool (e.g., swivel-360 [37], ChairIO [31] or different versions of the NaviChair [32], [37], [41], [42]) to control simulated self-motions in VR and are thus largely independent of the user's head position in space. For the HeadJoystick interface, however, we chose to track the user's head and use its position change to control simulated self-motions in VR for a number of reasons: Pre-tests showed that head movements seem to require less effort and are more precisely controllable than trunk movements, weight shifting, or chair/stool tilting, especially for smaller deflections. We hypothesized that this would contribute to overall usability, and support longer-term usage.

HeadJoystick also gives users the option to include as much or little upper-body movements and weight-shifting as they preferred and fit their body type and movement abilities. Finally, using head-tracking to control self-motion ensures that users always receive appropriate vestibular motion cueing signals in the direction of the virtual self-motion. We hypothesize that this helps to reduce visual-vestibular cue conflicts and in turn likely also motion sickness [17], [18].

As for the potential reasons for lower usability aspects of thumbsticks compared to HeadJoystick translation, in the post-experiment interview P1 said *"It [HeadJoystick] was intuitive with my body movements."* and P13 stated *"HeadJoystick was my favorite interface, because it was easy to use and learn."* Conversely, six participants mentioned that it was not easy to control 3 translational DoFs using a gamepad. For example, P4 said *"Gamepad was the worst interface, because its hard to control the movement. You can't go toward different directions easily."* We suggest that the Gamepad design may have contributed to its disadvantages compared to the embodied interfaces: While the mapping between input and the simulated motion matches for the head-based translation, gamepad or RC controllers usually split the four DoFs between two hands/thumbs, and mapping between input and the simulated motion does not match for all DoF. For example, it might not be intuitive to control simulated up/down translation and yaw rotation using a thumbstick pitch/roll rotation. Unfamiliarity of participants with our controller scheme of using left thumbstick for forward/backward and left/right motion and using right thumbstick for elevation and yaw rotation might be another potential reason for the lower performance and user ratings, even though no participants in pilot-tests or in the post-experiment interview mentioned such a barrier when using gamepad.

4.2 RQ2: Leaning-based interfaces improve flight performance

Results confirmed our tentative prediction about higher performance of leaning-based interfaces compared to thumb-

sticks (see Figure 3). HeadJoystick and LeaningTranslation seem to be the first 4DoF leaning-based flying interfaces that outperformed the prevalent and often highly familiar dual-thumbstick handheld interfaces both for short-term (Study 1) and repeated usage of the interface (Study 2).

Note that prior work on 4DoF leaning-based flying interfaces (e.g., head-rotation and head-translation in [23]) showed no performance benefit over hand-held interfaces, which could be due to the lower precision caused by the incorrect calibration, drifting from the zero-point, and the technical issues of having an actual drone. Prior research on 2DoF leaning-based flying interfaces showed mixed results regarding this research question: while the FlyJacket interface [62] did not improve performance, torso-strategy [5] improved only efficiency (more passed way-points). However, our results showed performance advantages of leaning-based flying interfaces in terms of not only efficiency (i.e., number of passed way-points) but also effectiveness including accuracy and precision.

Prior research on 2D (ground-based) leaning-based interfaces also showed mixed results in terms of improving navigation performance. For example, compared to 2D handheld interfaces, the majority of 2D leaning-based interfaces did not improve navigation performance (e.g., [31], [32], [34], [37], [42], [46]) except a few recent leaning-based interfaces (e.g., [40]) that reported improved performance in a navigational search task. Therefore, locomotion interface design guidelines usually suggested that 2D leaning-based interfaces provide reduced performance (e.g., [64]). Together with results from [40], our findings suggest that leaning-based interfaces indeed have the potential to outperform standard hand-held controller-based locomotion interfaces in both efficiency and effectiveness if designed well, for not only 2D (ground-based) but also 3D (flying), even when all 4DoF need to be controlled. Unlike ground-based leaning-based interfaces, our flying leaning-based interfaces showed higher accuracy/precision compared to the handheld interfaces, which could be due to controlling additional DoFs, which could increase complexity and thus reduce the navigation accuracy/precision when using handheld interfaces.

To explore the potential reasons for poor performance of the gamepad vs head-based translation, in the post-experiment interview, five participants mentioned that the gamepad was too sensitive for the later narrow tunnels compared to using the head. For example, P17 said *"it [gamepad] was too sensitive and I could not go easily to the narrow tunnels."* Lower movement range of thumbstick versus HeadJoystick could be a potential reason for the higher accuracy/precision of the head-based over thumbstick translation, as it might not be easy to fly with extremely low velocity when using thumbstick.

4.3 RQ3: Combining full physical rotation and leaning-based translation cues reduce motion sickness

Even though we limited the exposure/trial duration in Study 1 to 90s intentionally to reduce motion sickness, HeadJoystick was the only interface that did not increase motion sickness (post-pre trial) and showed significantly lower motion sickness than the gamepad. This implies that while providing both rotational and translational physical

self-motion cues can reduce motion sickness, neither of them alone might be enough to reduce motion sickness significantly. Our findings corroborate previous studies (e.g., [62]) that reported that FlyJacket 2DoF leaning-based flying interface reduced motion sickness compared to a hand-held interface. In the post-experiment interview, P24 said *"Gamepad is so difficult to use and with the highest level of sickness."* Further research is warranted to more closely assess how translation and rotation cues interact and contribute to motion sickness.

4.4 RQ4: Contributions of full physical rotations vs. leaning-based translations

Results confirmed our prediction that **embodied (leaning-based) translation** should improve user experience and performance compared to thumbstick translation, by showing significant benefits for all DV apart from motion sickness in both Study 1 and 2, including user experience factors, usability aspects, and performance measures (see Table 3 and Table 4). These findings are noteworthy as other promising leaning-based flying interfaces improved only a few user experience aspects (e.g., FlyJacket [62]) or one performance measure (e.g., torso-strategy [5]). The observed advantages of leaning-based translation could be useful for improving locomotion interfaces in situations where users have no access to a swivel chair or simply prefer not to rotate physically, e.g., due to convenience or laziness [78]. For example, when the user is sitting on a couch or non-rotating chair, or when using a stationary display like a TV or projection screen instead of an HMD. LeaningTranslation (without physical rotation) in Study 1 showed significant benefits over the gamepad for nine out of 15 measures and was the most favoured interfaces for four (out of 24) participants, who preferred rotating with the gamepad (instead of a chair). As an example, P4 stated *"LeaningTranslation was my favorite interface, because rotating with controller is easier."*

Results also confirmed our prediction that **embodied (physical) rotation** should improve user experience and performance compared to gamepad rotation, by showing significant benefits (main effects) in seven out of 12 DVs in Study 1. This clear benefit of physical rotations could also be relevant from the applied perspective, as most of the recent leaning-based flying interfaces did not allow for physical yaw rotation, including Dragon-riding [24], torso-strategy [5], FlyJacket [62], Head-Rotation, and Head-Translation [23], although there are a few exceptions (e.g., modified Flying-Head [23]). Thus, the observed clear advantages of leaning-based interfaces when using 1:1 360° physical rotation suggest that flying interface designers might want to consider allowing for full physical rotation to improve the overall user experience and performance.

The interaction between embodied (head-based) translation and embodied (physical) rotation suggests that combining embodied translation with embodied rotation can make the interface easier to use, easier to learn, and reduce task load. These findings could also help to understand why prior work reported inconsistent results regarding the impact of full physical rotation on 2D (ground-based) navigation (e.g., [52], [53]). Our results showed that the advantage of physical rotation depends on which translation technique

it is combined with. For example, when using gamepad translation, switching from gamepad to physical rotation improved none of the 15 measures. However, when using head-based translation, switching from gamepad to physical rotation not only improved enjoyment and preference ratings, but also revealed significant improvements in terms of ease of use, ease of learning, task load, long-term use, overall usability, and motion sickness. These results suggest that full physical rotation might improve the overall user experience only if it is combined with a suitable embodied translation technique, in the sense that both rotations and translations need to be embodied. This notion is corroborated by five participants mentioned in the post-experimental interview that controlling virtual translation with thumbs and virtual rotation with the body was confusing. For example, P2 explained that *"the worst interface was RealRotation, because it needs too much focus, both on your body and the gamepad."* These findings are aligned with prior concerns when using physical rotations with controller-based translations in 2D (ground-based) navigation such as [37] or informal observations of Grechkin and Riecke, [79].

The importance of providing both embodied rotation and translation is corroborated by post-experimental interview feedback: Nine participants mentioned that controlling both simulated translation and rotation using their body (instead of their hands) was more similar to real-world movement inside an actual spaceship rather than a game. E.g., P1 state that *"HeadJoystick was my favorite interface, because I don't need to think which part to control with my head and which part to control with my hand."* Embodied control of both simulated translation and rotation could be a potential reason for the usability advantages of the HeadJoystick in terms ease of use, ease of learning, and the task load, and might be related to an improved affordance [80]. Moreover, embodied locomotion frees up users' hands so they can use them for interaction with the environment, which has been stated as another advantage of hands-free locomotion by prior research on 2D navigation [42], [43], [44], [45]. As we found no prior research on the contributions of embodied translation with/without embodied rotation on user experience or performance for flying, all our findings in this regard expand the knowledge by addressing this gap.

4.5 RQ5: Leaning-based interfaces retain improved user experience, usability, and performance over repeated usage

Study 2 confirmed our hypothesis that the benefits of a leaning-based interface over a handheld interface in terms of user experience, usability, and performance observed in Study 1 will continue to hold even after extended (repeated) usage. Similar to prior studies (e.g., [34], [65], [68]), our study showed improved performance of leaning-based interfaces over repeated usage. However, unlike these prior works [34], [68], our second study showed that leaning-based interfaces such as HeadJoystick could have a faster performance improvement compared to using thumbsticks. That is, while both Controller and HeadJoystick showed significant learning effects in Study 2, performance improvements were more pronounced for the HeadJoystick: Even though during the first trial participants passed already

more than twice as many tunnels with the HeadJoystick than the Controller ($p < 0.001$), the subsequent learning effect and performance improvements were more pronounced for the HeadJoystick, indicated by the significant interaction between interface and trial number, and the steeper linear regression slope for the HeadJoystick (see Figure 4). Furthermore, linear regressions showed significant reductions of distance errors over the eight trials for the HeadJoystick, but not Controller. That is, even though participants were not familiar with the HeadJoystick, they already performed better with it in the first trials, and showed more pronounced improvements over time (as might be expected for novel interfaces) suggesting the full potential of leaning-based interfaces might be more apparent when allowing users sufficient practice.

While most of the measures for HeadJoystick and RealRotation/Controller were fairly similar between Study 1 and 2, extended usage in Study 2 showed more pronounced benefits of the HeadJoystick over RealRotation/Controller in terms of ease of learning, vection intensity, and task load. This might be related to the HeadJoystick being a novel interface for all participants and thus requiring more practice to reveal its full potential. That is, having sufficient time to learn the novel leaning-based interface and more intuitive control might allow users to more easily focus on their task and feel stronger vection, as they are less distracted by fiddling with the locomotion controls.

Study 2 showed a significant increase of motion sickness over the eight trials, similarly for both HeadJoystick and RealRotation/Controller. However, motion sickness overall remained fairly low ($< 6\%$) or < 28 for the total SSQ score, indicating that both interfaces are suitable for extended usage. The overall low motion sickness despite the fast-paced task and longer exposure in Study 2 suggests that the overall locomotion interfaces design and motion sickness mitigation measures of reducing maximum velocity, smoothing accelerations, and including embodied motion cues and thus reducing visual-vestibular cue conflicts were suitable, and can help guide future interface designs.

4.6 Limitations

While results from Study 1 and 2 are fairly consistent and show overall substantial effects and effect sizes, there are several potential limitations that could guide future research: Although Study 2 corroborated and largely replicated findings from Study 1 for repeated (extended) usage, we only ran 8 trials of 60s per interface. Future research is needed to investigate if/how our findings might extend to much longer durations or usage across several days/weeks, which can be relevant for real-world applications. All participants were familiar with the gamepad (but not the HeadJoystick), which could have affected our results too. As we designed this drone-racing task without using actual drones due to the high chance of colliding with the narrow tunnels, future research will need to test how the results generalize to telepresence applications with actual quadcopter drones.

5 CONCLUSION AND FUTURE WORK

In this paper, we introduced HeadJoystick, a novel 4DoF leaning-based flying interface for VR applications. In pre-

vious work, leaning-based flying interfaces for 2DoF flying improved either user experience aspects (e.g., FlyJacket [62]) or performance (e.g., torso-strategy [5]), but not both. In contrast, we showed that compared to handheld flying interfaces, HeadJoystick improved six user experience factors (i.e., enjoyment, taskload, immersion, presence, Vection intensity, and overall preference), six usability aspects (i.e., motion sickness, ease of use, ease of learning, long-term use, daily use, and overall usability), and three performance measures (i.e., efficiency, precision, and accuracy). We did so in a VR-simulated drone waypoint navigation task. In addition, we corroborated these benefits under repeated exposure, with improved performance and only minimal increases in motion sickness over time. Together, this provides promising first evidence that leaning-based interfaces can improve performance and usability/user experience not just for 2DoF (fixed-wing) flight [5], [62], but also in 4DoF flying (similar to quadcopter drones). Our results could also benefit telepresence applications as they share similar challenges of using handheld controllers, even though we did not specifically investigate those.

From an applied perspective, HeadJoystick is easy to set up and affordable as it requires no additional hardware besides a swivel chair commonly found in most homes and offices, thus can be readily integrated into existing VR setups that provide 6DoF tracking. Although we only tested HeadJoystick with seated users, it can be easily adapted to standing, and pilot tests were promising. In applications where HeadJoystick could be used for tasks that require free body movements (such as conversation with a fellow visitor during virtual tourism), an "activate" switch for HeadJoystick could be considered, so users can choose to be completely stationary and move their body freely whenever they do not plan to locomote.

In situations where physical user rotation is not desired (e.g., due to convenience or laziness [78]) or feasible (e.g., when sitting on a couch or on transit/planes, or using a projection or TV screen instead of a HMD), using LeaningTranslation provides considerable advantages over a gamepad in terms of six user experience measures as well as three performance measures. Compared to HeadJoystick and LeaningTranslation, other promising leaning-based flying interfaces (e.g., torso-strategy [5] and FlyJacket [62]) might not be as suitable for daily real-life applications. For example, torso-strategy requires attaching several camera-based motion tracker markers to the user's upper-body to measure the flexion/extension of the trunk muscles during flight [5], and FlyJacket requires the user to wear a backpack, which holds his/her arms up during flight [62].

Future research is needed to investigate how the current findings and advantages observed for Head-Joystick and Leaning-translation might generalize to different virtual or telepresence tasks such as 2D navigation, driving, navigation with secondary interaction task (e.g., First-person shooter games), and 3D telepresence scenarios with quadcopter drones using RC controllers. Future studies can also investigate our suggested interfaces in more detail, such as standing as compared to sitting users [81], and more diverse participant samples. Overall, these findings extend our knowledge about the advantages of the leaning-based flying interfaces in general and specifically our suggested

interface, HeadJoystick, as well as the contributions and interactions of embodied rotational versus translational cues.

REFERENCES

- [1] P. U. Lima, "Search and rescue robots: The civil protection teams of the future," in *Third International Conference on IEEE Emerging Security Technologies*, 2012, pp. 12–19.
- [2] D. Mirk and H. Hlavacs, "Virtual Tourism with Drones: Experiments and Lag Compensation," in *Proceedings of the First Workshop on Micro Aerial Vehicle Networks, Systems, and Applications for Civilian Use*, ser. DroNet '15. NY, USA: ACM, 2015, pp. 45–50.
- [3] E. R. Stepanova, M. von der Heyde, A. Kitson, T. Schiphorst, and B. E. Riecke, "Gathering and applying guidelines for mobile robot design for urban search and rescue application," in *International Conference on Human-Computer Interaction*. Springer, 2017, pp. 562–581.
- [4] D. A. Bowman, E. Kruijff, J. J. LaViola, and I. Poupyrev, *3D User Interfaces: Theory and Practice*. Addison-Wesley Professional, 2017.
- [5] J. Miehlsbradt, A. Cherpillod, S. Mintchev, M. Coscia, F. Artoni, D. Floreano, and S. Micera, "Data-driven body-machine interface for the accurate control of drones," *Proceedings of the National Academy of Sciences*, pp. 7913–7918, Jul. 2018.
- [6] M. Quigley, M. A. Goodrich, and R. W. Beard, "Semi-autonomous human-UAV interfaces for fixed-wing mini-UAVs," in *IEEE/RSJ IROS*, vol. 3, 2004, pp. 2457–2462.
- [7] J. R. Cauchard, K. Y. Zhai, and J. A. Landay, "Drone & me: an exploration into natural human-drone interaction," in *Proceedings of the ACM international joint conference on pervasive and ubiquitous computing*. ACM, 2015, pp. 361–365.
- [8] K. Ikeuchi, T. Otsuka, A. Yoshii, M. Sakamoto, and T. Nakajima, "KinecDrone: enhancing somatic sensation to fly in the sky with Kinect and AR. Drone," in *Proceedings of the 5th Augmented Human International Conference*. ACM, 2014, p. 53.
- [9] M. Monajjemi, S. Mohaimenianpour, and R. Vaughan, "UAV, come to me: End-to-end, multi-scale situated HRI with an uninstrumented human and a distant UAV," in *2016 IEEE/RSJ IROS*, 2016, pp. 4410–4417.
- [10] K. Pfeil, S. L. Koh, and J. LaViola, "Exploring 3d gesture metaphors for interaction with unmanned aerial vehicles," in *Proceedings of the international conference on Intelligent user interfaces*. ACM, 2013, pp. 257–266.
- [11] A. Sarkar, K. A. Patel, R. G. Ram, and G. K. Kapoor, "Gesture control of drone using a motion controller," in *International Conference on Industrial Informatics and Computer Systems*. IEEE, 2016, pp. 1–5.
- [12] A. Stoica, F. Salvioli, and C. Flowers, "Remote Control of Quadrotor Teams, Using Hand Gestures." 9th ACM/IEEE HRI, 2014, pp. 296–297.
- [13] R. V. Krishna, B. S. Sathish, P. Ganesan, P. J. Babu, and R. Abilash, "Design of voice and gesture controlled Quadcopter," in *2015 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS)*, Mar. 2015, pp. 1–6.
- [14] E. Peshkova, M. Hitz, and D. Ahlström, "Exploring user-defined gestures and voice commands to control an unmanned aerial vehicle," in *International Conference on Intelligent Technologies for Interactive Entertainment*. Springer, 2016, pp. 47–62.
- [15] Y. Yu, D. He, W. Hua, S. Li, Y. Qi, Y. Wang, and G. Pan, "FlyingBuddy2: a brain-controlled assistant for the handicapped." in *Ubicomp*, 2012, pp. 669–670.
- [16] B. D. Lawson and B. E. Riecke, "Perception of Body Motion." in *Hale KS, Stanney KM (eds) Handbook of Virtual Environments: Design, Implementation, and Applications*, 2nd ed. CRC Press, 2014, pp. 163–195.
- [17] J. T. Reason and J. J. Brand, *Motion sickness*. Oxford, England: Academic Press, 1975, vol. vii.
- [18] R. S. Kennedy, J. Drexler, and R. C. Kennedy, "Research in visually induced motion sickness," *Applied ergonomics*, vol. 41, no. 4, pp. 494–503, 2010.
- [19] J. B. van Erp, M. Duistermaat, C. Jansen, E. Groen, and M. Hoedemaeker, "Tele-presence: Bringing the operator back in the loop," 2006, pp. 9–1 – 9–18.
- [20] B. E. Riecke and J. D. Jordan, "Comparing the effectiveness of different displays in enhancing illusions of self-movement (vection)," *Frontiers in psychology*, vol. 6, p. 713, 2015.
- [21] E. L. Groen and W. Bles, "How to use body tilt for the simulation of linear self motion," *Journal of Vestibular Research: Equilibrium & Orientation*, vol. 14, no. 5, pp. 375–385, Sep. 2004.

- [22] P. Miermeister, M. Lächele, R. Boss, C. Masone, C. Schenk, J. Tesch, M. Kerger, H. Teufel, A. Pott, and H. H. Bühlhoff, "The cablerobot simulator large scale motion platform based on cable robot technology," in *2016 IEEE/RSJ IROS*, 2016, pp. 3024–3029.
- [23] C. Pittman and J. J. LaViola, Jr., "Exploring Head Tracked Head Mounted Displays for First Person Robot Teleoperation," in *Proceedings of the 19th International Conference on Intelligent User Interfaces*, ser. IUI '14. NY, USA: ACM, 2014, pp. 323–328.
- [24] B. Schulte, C. P. Bimberg, J. Hertel, F. Neu, P. Heidenreich, G. Bruder, and P. Lubos, "Evaluation of Two Leaning-Based Control Methods in a Dragon Riding Interface," in *Proceedings of the GI-Workshop VR/AR*, 2016, pp. 109–116.
- [25] D. Krupke, P. Lubos, L. Demski, J. Brinkhoff, G. Weber, F. Willke, and F. Steinicke, "Control methods in a supernatural flight simulator," in *IEEE VR*, Mar. 2016, pp. 329–329.
- [26] M. Perusquía-Hernández, T. Enomoto, T. Martins, M. Otsuki, H. Iwata, and K. Suzuki, "Embodied interface for levitation and navigation in a 3d large space." 8th Augmented Human International Conference, 2017, p. 4.
- [27] D. Krupke, P. Lubos, L. Demski, J. Brinkhoff, G. Weber, F. Willke, and F. Steinicke, "Evaluation of Control Methods in a Supernatural Zero-Gravity Flight Simulator," in *Proceedings of the GI-Workshop VR/AR*, 2015.
- [28] E. Viirre, B. J. Price, and B. Chase, "direct effects of virtual environments on users," in *K. Hale and K. Stanney, Handbook of Virtual Environments*, 2nd ed. CRC Press, 2015, pp. 521–529.
- [29] M. Rheiner, "Birdly an Attempt to Fly," in *ACM SIGGRAPH Emerging Technologies*. NY, USA: ACM, 2014, pp. 3:1–3:1.
- [30] A. Cherpillod, S. Mintchev, and D. Floreano, "Embodied Flight with a Drone," *Third IEEE IRC*, pp. 386–390, Jul. 2017.
- [31] S. Beckhaus, K. J. Blom, and M. Haringer, "A new gaming device and interaction method for a First-Person-Shooter," *Proceedings of the Computer Science and Magic*, 2005.
- [32] A. Kitson, A. M. Hashemian, K. Stepanova, and B. E. Riecke, "Comparing Leaning-Based Motion Cueing Interfaces for Virtual Reality Locomotion." *IEEE 3DUI*, 2017, pp. 73–82.
- [33] B. E. Riecke and D. Feureissen, "To Move or Not to Move: Can Active Control and User-driven Motion Cueing Enhance Self-motion Perception ("Vection") in Virtual Reality?" in *Proceedings of the ACM SAP*. NY, USA: ACM, 2012, pp. 17–24.
- [34] M. Marchal, J. Pettré, and A. Lécuyer, "Joyman: A human-scale joystick for navigating in virtual worlds." *IEEE 3DUI*, Mar. 2011, pp. 19–26.
- [35] A. Harris, K. Nguyen, P. T. Wilson, M. Jackoski, and B. Williams, "Human Joystick: Wii-leaning to Translate in Large Virtual Environments," in *Proceedings of the 13th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and Its Applications in Industry*. NY, USA: ACM, 2014, pp. 231–234.
- [36] E. Kruijff, A. Marquardt, C. Trepkowski, R. W. Lindeman, A. Hinkenjann, J. Maiero, and B. Riecke, "On Your Feet! Enhancing Self-Motion Perception in Leaning-Based Interfaces through Multisensory Stimuli." *ACM SUI*, Tokyo, Japan, 2016.
- [37] A. M. Hashemian and B. E. Riecke, "Swivel-Chair: Evaluating Seated Full-Rotational Interfaces for Virtual Reality Navigation." Vancouver, Canada: Software for Augmented and Virtual Reality (SAVR), Oct. 2017.
- [38] B. E. Riecke, "Simple user-generated motion cueing can enhance self-motion perception (vection) in virtual reality," in *Proceedings of the ACM symposium on Virtual reality software and technology*. Limassol, Cyprus: ACM, 2006, pp. 104–107.
- [39] B. E. Riecke, D. Feureissen, and J. Rieser, "Auditory self-motion illusions ("circular vection") can be facilitated by vibrations and the potential for actual motion." NY, USA: ACM Press, 2008, pp. 147–154.
- [40] T. Nguyen-Vo, B. E. Riecke, W. Stuerzlinger, D.-M. Pham, and E. Kruijff, "NaviBoard and NaviChair: Limited Translation Combined with Full Rotation for Efficient Virtual Locomotion," *IEEE TVCG*, 2019.
- [41] J. Freiberg, "Experience Before Construction: Immersive Virtual Reality Design Tools for Architectural Practice." Master's thesis, Simon Fraser University, Surrey, BC, Canada., 2015.
- [42] A. Kitson, B. E. Riecke, A. M. Hashemian, and C. Neustaedter, "NaviChair: Evaluating an Embodied Interface Using a Pointing Task to Navigate Virtual Reality," in *Proceedings of the 3rd ACM SUI*, NY, USA, 2015, pp. 123–126.
- [43] S. Beckhaus, K. J. Blom, and M. Haringer, "Intuitive, hands-free travel interfaces for virtual environments," in *New Directions in 3D User Interfaces Workshop of IEEE VR*, 2005, pp. 57–60.
- [44] J. J. LaViola, Jr., D. A. Feliz, D. F. Keefe, and R. C. Zeleznik, "Hands-free Multi-scale Navigation in Virtual Environments," in *Symposium on I3D*. NY, USA: ACM, 2001, pp. 9–15.
- [45] D. Zielasko, S. Horn, S. Freitag, B. Weyers, and T. W. Kuhlen, "Evaluation of hands-free HMD-based navigation techniques for immersive data analysis." *IEEE 3DUI*, Mar. 2016, pp. 113–119.
- [46] R. P. McMahan, D. A. Bowman, D. J. Zielinski, and R. B. Bardy, "Evaluating Display Fidelity and Interaction Fidelity in a Virtual Reality Game," *IEEE TVCG*, vol. 18, no. 4, pp. 626–633, 2012.
- [47] M. J. Farrell and I. H. Robertson, "Mental rotation and automatic updating of body-centered spatial relationships." *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 24, no. 1, p. 227, 1998.
- [48] C. C. Presson and D. R. Montello, "Updating after rotational and translational body movements: Coordinate structure of perspective space," *Perception*, vol. 23, no. 12, pp. 1447–1455, 1994.
- [49] J. J. Rieser, "Access to knowledge of spatial structure at novel points of observation," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 15, no. 6, pp. 1157–1165, 1989.
- [50] R. L. Klatzky, J. M. Loomis, A. C. Beall, S. S. Chance, and R. G. Golledge, "Spatial updating of self-position and orientation during real, imagined, and virtual locomotion," *Psychological science*, vol. 9, no. 4, pp. 293–298, 1998.
- [51] R. A. Ruddle, S. J. Payne, and D. M. Jones, "Navigating large-scale virtual environments: what differences occur between helmet-mounted and desk-top displays?" *Presence: Teleoperators & Virtual Environments*, vol. 8, no. 2, pp. 157–168, 1999, publisher: MIT Press.
- [52] R. A. Ruddle, "The Effect of Translational and Rotational Body-Based Information on Navigation," in *Human Walking in Virtual Environments*, F. Steinicke, Y. Visell, J. Campos, and A. Lécuyer, Eds. Springer New York, 2013, pp. 99–112.
- [53] B. E. Riecke, B. Bodenheimer, T. P. McNamara, B. Williams, P. Peng, and D. Feureissen, "Do we need to walk for effective virtual reality navigation? physical rotations alone may suffice," in *International Conference on Spatial Cognition VII*. Springer, 2010, pp. 234–247.
- [54] E. Guy, P. Punpongson, D. Iwai, K. Sato, and T. Boubekeur, "LazyNav: 3D ground navigation with non-critical body parts," in *IEEE Symposium on 3D User Interfaces (3DUI)*, Mar. 2015, pp. 43–50.
- [55] E. Langbehn, T. Eichler, S. Ghose, K. von Luck, G. Bruder, and F. Steinicke, "Evaluation of an Omnidirectional Walking-in-Place User Interface with Virtual Locomotion Speed Scaled by Forward Leaning Angle," in *Proceedings of the GI Workshop on Virtual and Augmented Reality (GI VR/AR)*, 2015, pp. 149–160.
- [56] C. Wang, O. A. Dogaru, P. L. Strandholt, N. C. Nilsson, R. Nordahl, and S. Serafin, "Step aside: an initial exploration of gestural input for lateral movement during walking-in-place locomotion," in *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology*, 2018, pp. 1–5.
- [57] M. G. Silva and D. A. Bowman, "Body-based Interaction for Desktop Games," in *CHI Extended Abstracts on Human Factors in Computing Systems*. NY, USA: ACM, 2009, pp. 4249–4254.
- [58] M. R. Chester, M. J. Rys, and S. A. Konz, "Leg swelling, comfort and fatigue when sitting, standing, and sit/standing," *International Journal of Industrial Ergonomics*, vol. 29, no. 5, pp. 289–296, 2002.
- [59] O. Merhi, E. Faugloire, M. Flanagan, and T. A. Stoffregen, "Motion sickness, console video games, and head-mounted displays," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 49, no. 5, pp. 920–934, 2007.
- [60] G. E. Riccio and T. A. Stoffregen, "An ecological theory of motion sickness and postural instability," *Ecological psychology*, vol. 3, no. 3, pp. 195–240, 1991.
- [61] D. R. Badcock, S. Palmisano, and J. G. May, "Vision and Virtual Environments," in *K. S. Hale and K. M. Stanney (Eds.), Handbook of Virtual Environments (2nd ed.)*. CRC Press, 2014, pp. 39–84.
- [62] C. Rognon, S. Mintchev, F. Dell'Agnola, A. Cherpillod, D. Atienza, and D. Floreano, "FlyJacket: An Upper Body Soft Exoskeleton for Immersive Drone Control," *IEEE Robotics and Automation Letters*, vol. 3, no. 3, pp. 2362–2369, Jul. 2018.
- [63] K. Higuchi, K. Fujii, and J. Rekimoto, "Flying head: A head-synchronization mechanism for flying telepresence." 23rd ICAT, Dec. 2013, pp. 28–34.

- [64] D. A. Bowman, R. P. McMahan, and E. D. Ragan, "Questioning naturalism in 3d user interfaces," *Communications of the ACM*, vol. 55, no. 9, pp. 78–88, 2012.
- [65] J. Wang and R. W. Lindeman, "Comparing isometric and elastic surfboard interfaces for leaning-based travel in 3D virtual environments," in *3D User Interfaces (3DUI), 2012 IEEE Symposium on*. IEEE, 2012, pp. 31–38.
- [66] A. M. Hashemian and B. E. Riecke, "Leaning-Based 360deg Interfaces: Investigating Virtual Reality Navigation Interfaces with Leaning-Based-Translation and Full-Rotation," in *Proceedings of HCI International Conference on Virtual, Augmented and Mixed Reality*, vol. 10280. Springer, Jul. 2017, pp. 15–32.
- [67] R. P. McMahan, A. J. D. Alon, S. Lazem, R. J. Beaton, D. Machaj, M. Schaefer, M. G. Silva, A. Leal, R. Hagan, and D. A. Bowman, "Evaluating natural interaction techniques in video games," in *2010 IEEE Symposium on 3D User Interfaces (3DUI)*. IEEE, 2010, pp. 11–14.
- [68] L. Terziman, M. Marchal, M. Emily, F. Multon, B. Arnaldi, and A. Lécuyer, "Shake-your-head: Revisiting Walking-in-place for Desktop Virtual Reality," in *Proceedings of the 17th ACM Symposium on Virtual Reality Software and Technology*, ser. VRST '10. New York, NY, USA: ACM, 2010, pp. 27–34.
- [69] B. D. Lawson, "Motion Sickness Symptomatology and Origins." in *K. Hale and K. Stanney (Eds.), Handbook of Virtual Environments (2nd ed., pp. 531–600)*. CRC Press, 2014, pp. 531–600.
- [70] D. R. Trindade and A. B. Raposo, "Improving 3d navigation techniques in multiscale environments: a cubemap-based approach," *Multimedia Tools and Applications*, vol. 73, no. 2, pp. 939–959, Nov. 2014.
- [71] J. Wang and R. Lindeman, "Leaning-based Travel Interfaces Revisited: Frontal Versus Sidewise Stances for Flying in 3d Virtual Spaces," in *Proceedings of the 18th ACM Symposium on Virtual Reality Software and Technology*. NY, USA: ACM, 2012, pp. 121–128.
- [72] E. Sikstrom, A. d. Gotzen, and S. Serafin, "Wings and flying in immersive VR #x2014; Controller type, sound effects and experienced ownership and agency." *IEEE VR*, Mar. 2015, pp. 281–282.
- [73] R. McMahan, R. Kopper, and D. Bowman, "Principles for Designing Effective 3D Interaction Techniques," in *K. Hale and K. Stanney, Handbook of Virtual Environments*. CRC Press, 2014, pp. 285–311.
- [74] S. C. de Vries and P. Padmos, "Remotely Controlled Flying Aided by a Head-Slaved Camera and HMD." *Human Factors Research Institute, Soesterberg (Netherlands)*, Tech. Rep., 1997.
- [75] M. Slater, J. McCarthy, and F. Maringelli, "The influence of body movement on subjective presence in virtual environments," *Human Factors*, vol. 40, no. 3, pp. 469–477, 1998.
- [76] S. G. Hart, "NASA-task load index (NASA-TLX); 20 years later," in *Proceedings of the human factors and ergonomics society annual meeting*, vol. 50. Sage Publications Sage CA: Los Angeles, CA, 2006, pp. 904–908.
- [77] 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, 1993.
- [78] E. D. Ragan, S. Scerbo, F. Bacim, and D. A. Bowman, "Amplified Head Rotation in Virtual Reality and the Effects on 3d Search, Training Transfer, and Spatial Orientation," *IEEE TVCG*, vol. 23, no. 8, pp. 1880–1895, Aug. 2017.
- [79] T. Y. Grechkin and B. E. Riecke, "Re-evaluating Benefits of Body-based Rotational Cues for Maintaining Orientation in Virtual Environments: Men Benefit from Real Rotations, Women Don'T," in *Proceedings of the ACM SAP*. NY, USA: ACM, 2014, pp. 99–102.
- [80] B. E. Riecke and D. Zielasko, "Towards an affordance of embodied locomotion interfaces in VR: How to know how to move?" in *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. Atlanta, GA, USA: IEEE, Mar. 2020, pp. 295–296.
- [81] D. Zielasko and B. E. Riecke, "Sitting vs. Standing in VR: Towards a Systematic Classification of Challenges and (Dis)Advantages," in *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. Atlanta, GA, USA: IEEE, Mar. 2020, pp. 297–298.



Abraham M. Hashemian Abraham has a MSc in Computer Engineering - Artificial Intelligence from Sharif University of Technology. He started his PhD at School of Interactive Arts and Technology (SIAT), Simon Fraser University (SFU) in Spring 2015, and joined the iSPACE research lab, where he started his studies on VR technology and locomotion in immersive virtual environments. His research interests involve using technologies such as virtual and augmented reality, artificial intelligence, and machine learning

in gaming to improve human life using games.



Matin Lotfaliei Matin is an experienced programmer with a background in Computer Science. During his BSc, he started following his area of interests in Human-Computer Interaction, Artificial Intelligence, and Virtual Reality. After graduation, he remotely collaborated with the iSPACE research lab in 2016. Eventually in Fall 2017, he joined the team as a MSc student at the School of Interactive Arts and Technology (SIAT), Simon Fraser University (SFU), where he studied embodied locomotion interfaces and visually induced motion sickness in virtual environments.

visually induced motion sickness in virtual environments.



Ashu Adhikari has a bachelor's degree in Electronics and Communication Engineering. He became engaged in virtual reality development during his senior year and has been following his interests in Human-Computer Interaction ever since. In Fall 2018, he enrolled in the School of Interactive Arts and Technology (SIAT), Simon Fraser University (SFU) as a MSc student and joined the iSPACE team. He pursues his interest in embodied locomotion interfaces in virtual reality.



Ernst Kruijff Ernst Kruijff is a Professor for Human Computer Interaction at the Institute of Visual Computing, Department of Computer Science, Bonn-Rhein-Sieg University of Applied Sciences. He is also adjunct professor at SFU-SIAT in Canada. For almost two decades, his research has focused at the human-factors driven analysis, design and validation of multisensory 3D user interfaces. His research looks at the usage of audio-tactile feedback methods to enhance interaction and perception within the frame of AR view management, VR navigation and hybrid 2D/3D mobile systems.

frame of AR view management, VR navigation and hybrid 2D/3D mobile systems.



Bernhard E. Riecke After researching for a decade in the VR Group of the Max Planck Institute in Tübingen, Germany and working as a post-doc at Vanderbilt and UC Santa Barbara, Dr. Riecke joined SFU-SIAT in 2008. His research combines multidisciplinary research approaches and VR/xR to investigate: Human spatial cognition/orientation/updates/navigation; Self-motion perception, illusions ("vection"), and simulation; Designing for transformative positive experiences, and rapid prototyping and agile researching in xR.

experiences, and rapid prototyping and agile researching in xR.