Versatile Mixed-method Locomotion under Free-hand and Controller-based Virtual Reality Interfaces

Anton Franzluebbers
Kyle Johnsen
anton@uga.edu
kjohnsen@uga.edu
Virtual Experiences Lab, University of Georgia
Athens, Georgia, USA

Figure 1: Three tasks completed under free-hand and controller-based interfaces within the *conVRged* social VR application.

ABSTRACT

Locomotion systems that allow the user to interact with large virtual spaces require precise input, competing with the same inputs available for performing a task in the virtual world. Despite extensive research on hand tracking input modalities, there is a lack of a widely adopted mechanism that offers general-purpose, high-precision locomotion across various applications. This research aims to address this gap by proposing a design that combines teleportation with a grab-pull locomotion scheme to bridge the divide between long-distance and high-precision locomotion in both a tracked-controller and free-hand environment. The implementation details for both tracked controller and tracked hand environments are presented and evaluated through a user study. The study findings indicate that each locomotion mechanism holds value for different tasks, with grab-pull providing more benefit in scenarios where smaller, more precise positioning is required. As found in prior research, controller tracking was found to be faster than hand tracking, but all participants were able to successfully use the locomotion system with both interfaces.

CCS CONCEPTS

• **Human-centered computing** → *Gestural input*; **Virtual reality**; **Mixed / augmented reality**; *Empirical studies in HCI*.

KEYWORDS

virtual reality, locomotion, free-hand, controller, hand tracking, interaction

ACM Reference Format:
Anton Franzluebbers and Kyle Johnsen. 2023. Versatile Mixed-method Locomotion under Free-hand and Controller-based Virtual Reality Interfaces. In *29th ACM Symposium on Virtual Reality Software and Technology (VRST 2023), October 09–11, 2023, Christchurch, New Zealand*. ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3611659.3615701

VRST 2023, October 09–11, 2023, Christchurch, New Zealand

ACM ISBN 979-8-4007-0328-7/23/10…$15.00
https://doi.org/10.1145/3611659.3615701

1 INTRODUCTION

Visual hand tracking is becoming a standard feature of Augmented and Virtual Reality (AR/VR) systems, with some systems offering an option for hand tracking or controller tracking (e.g., Meta Quest 1/2/Pro, Pico Neo 4) and others relying on hand tracking exclusively (e.g., Microsoft Hololens, upcoming Apple Vision Pro). Eliminating the requirement of controllers is attractive as hands offer a more convenient, always available, and perhaps more natural interface. However, since hands have neither buttons nor tracking aids such as LEDs for more precise positioning, there are considerable challenges in user interface design. This paper discusses how we designed such a free-hand interface, attempting to maintain the versatility of a controller-based interface to support many complex tasks and to keep usability high enough to warrant its use.

A key aspect of this work that differentiates it from other studies on free-hand interfaces is that it is a study built for, *and within*, a real-world social VR application. That application is currently used for a number of other projects, and has an established and versatile controller-based interface already. The free-hand interface was intended as a convenient option for when controllers are not available (e.g., dead batteries) or not desirable, with the understanding that it would likely be less usable due to intrinsic issues in hand tracking (e.g., lack of buttons) and technical issues in tracking the hands with head-mounted cameras. Thus, the primary purpose of the current work was not to build an equivalently usable system, but instead to minimize and then quantify the usability differences.

We especially focused on locomotion in the design of the interface. The controller-based interface provides teleporting for large distance movements, fixed increment turning, and a grab-pull metaphor for precise positioning, relying on multiple controls for activation and excellent controller position and orientation tracking for control. Providing the same movement versatility in the hand-based locomotion interface required extensive adaptation. Instead of multiple buttons, a single index-thumb pinch gesture was chosen because it was one of the more reliably detected hand gestures and is commonly used in the Meta Quest home interface, increasing familiarity. Together with the pinch gesture, region-based, contextual activation with visual feedback resolves ambiguities and minimizes errors, and filtering is employed to increase accuracy.

After finalizing the design of the interface, we conducted a user study with 20 participants where we gathered detailed performance data while they completed a variety of tasks with the controller-based and free-hand interfaces. In addition to raw performance statistics such as travel time, we were interested in seeing how participants leveraged the various techniques when they had the option. Results from the study, including user feedback, confirmed our expectations that controllers were significantly faster for performing nearly all tasks. However, all users were able to complete all tasks, without reported complaints about simulator sickness and only two reported instances of fatigue. This suggests that the free-hand is a worthwhile addition, one that is likely to continue to improve as hand-tracking quality increases.

2 RELATED WORKS

There is a wide breadth of research that evaluates locomotion techniques in VR, but only a subset is focused on how they work with hand tracking, though several compare hand tracking to controller tracking for other tasks [10, 15]. Schäfer et al. evaluated four different teleportation techniques designed for use within a hand-tracked environment [21]. Two of these were two-handed techniques, which used a gesture on the other hand to activate the teleport; and two were one-handed, which used a "dwell" technique to activate the teleport. The evaluation task was a long hallway with "primitive graphics style" with 10 columns at random intervals for teleport targets. The authors found no significant advantage for two-handed versus one-handed gestures, concluding that one-handed interaction was sufficient for a usable teleporting interface. However while teleporting was disabled while the hands were near an interactable object, it is not clear whether this the ability to activate the teleporter ray using the tested gestures would interfere with more complex task interaction or social gestures. Work by Zielasko et al. also compared controller-free locomotion methods to a traditional controller, recommending methods with more freedom of movement such as the ability to move backward, though the controller-free methods did not use tracked hand position as the input method for locomotion [28].

Kim et al. explored a variety of non-controller input methods including hand-tracking, eye-tracking, and electroencephalography (EEG) [11]. A user study found "relative superiority of eye-tracking for location targeting and hand-tracking for teleport triggering." Our own study took inspiration from the teleporting task design, which consisted of a hallway with teleport targets on alternating sides. The spacing of the targets was such that the angle to the next target included a variety of angles and distances, and the alternating design ensured even counts of left and right turns. Both of these two works used an external Leap Motion module for their hand tracking. Visual tracking of hands has only become popular in recent decades, but other systems have used gloves to more accurately track gestures such as the work from Mapes et al. in 1995 [13].

Outside of teleporting, some works look at the use of hand tracking to control the speed or direction of smooth locomotion techniques. Zhao et al. used unique hand gestures for locomotion, including a "finger distance" gesture where the distance between the index finger an the thumb of a hand would determine the movement speed of the locomotion system, or another where the number of fingers displayed would increment the movement speed [27]. Their third system used the frequency of finger tapping for speed control. These systems were compared to an untracked gamepad in a user study performed outside of immersive VR.

Though it is relatively common in popular VR games (e.g. Nock[1], Lone Echo/Echo VR[2], Gorn[3], Brass Tactics[4], Google Earth VR[5], and to a lesser extent Gorilla Tag[6], etc.) hand-based grab-to-move locomotion has a relatively small body of published research [6]. Arm-swinging could be seen as a related technique, but it generally is a mechanism to influence the velocity of a sliding locomotion technique rather than direct manipulation over the user's position in 3 degrees of freedom [4, 17, 24]. Rantala et al. included a grab-pull system in their evaluation of several locomotion techniques for a visual observation task, and found that it outperformed a non-continuous teleport interface for their task, which involved counting small dots along a hallway as they moved [19]. Lim et

[1]https://nock.game/
[2]http://echo.games/
[3]https://store.steampowered.com/app/578620/GORN/
[4]https://www.oculus.com/experiences/rift/1101975213197949/
[5]https://vr.google.com/earth/
[6]https://www.gorillatagvr.com/

al. also compared three locomotion techniques that included grab-pull, but focused on the ability for this technique to move in three dimensions [12].

In addition to locomotion, grabbing and interacting with objects in the virtual world must be designed for a high quality multi-purpose VR application. Schäfer et al. compared a controller to a hand-tracked interface for picking up and placing objects [20]. They found significantly better performance when using a tracked controller, but did not find any significant differences between their proposed hand tracking grab techniques, either a forefinger pinch or a whole-hand grab. Masurovsky et al. and Hameed et al performed similar hand-controller comparisons, and both found significantly better performance when using tracked controllers compared to controller-free interactions [8, 14].

Other works tried to combine the benefits of both controller and controller-free inputs. Capece et al. presented a system that used a tracked controller in one hand for the use of the precise button inputs, and pinch with hand tracking to interact with the nodes on a 3D graph [5]. All locomotion inputs were performed using directional input on the controller.

3 SYSTEM DESIGN

3.1 Environment

The work presented here is built in the context of *conVRged*, a multi-purpose social VR world developed in our lab across several research and outreach projects, which is in the process of being converted to a fully open source project[7]. The application is designed such that it can contain a variety of experiences, from presentation-focused environments with screen-sharing, shared whiteboards, to games such as chess, checkers, remote-controlled cars, and a playground. The system consists of several separate environments for each target application. An elevator metaphor is used to travel between environments, with the buttons in the elevator used to select a target environment. While *conVRged* is used within our research group, its design mimics aspects of several commercial applications, such as VR Chat, Rec Room, or Horizon Workrooms.

While the application supports custom avatars (via Ready Player Me[8]), for the study we opted to use an androgynous avatar model with a pure white texture, as seen in Figure 1. This avatar was present whether the user is using hand tracking or traditional controller tracking, but instead of using the hand models provided by Ready Player Me for hand tracking on the Quest 2, the standard Meta hand models were used, which match the size and shape of the user's real hands.

3.2 Interaction

Most of the environment-specific interaction in the application is accomplished by a grabbing metaphor, wherein any movable object can be temporarily attached to the virtual hand and then released or thrown. This is accomplished by moving the avatar hand close to an object, which highlights if it can be picked up, and then activating and deactivating a grab intent. When using the Quest 2 controllers, this action is performed with either hand-trigger. When using hand-tracking, it is performed by bringing the tips of the index and thumb finger close enough together (pinching). Additionally, the avatar hands can be used on virtual touch screens by intersecting the virtual hand finger tip with a control or by raycasting from a distance.

[7]https://github.com/velaboratory/conVRged-OSS
[8]https://readyplayer.me

3.3 Locomotion

The locomotion system complements physical user movement (1-1 with real world movement) with virtual teleportation and grab-based sliding in any direction. The combination of these supports diverse physical and virtual environments and tasks.

3.3.1 *Grab-pull.*

Some tasks, like chess, are performed on a very small scale. In the real world, chess players are generally seated in front of the board, and do not get up or move around much. However, the initial body placement can be difficult, especially if seated, and depends on the height and comfortable reach of the person. In order to enable a comfortable playing position in VR, precise micro-adjustments must be possible beyond that of a point and teleport mechanism. To achieve this, a "grab-pull" sliding system was implemented. The user can use a button on the controller or a pinch activation gesture in hand tracking mode to lock their hand to the world, then when they move their hand, their body would move to keep their hand in the same location. Momentum is preserved when releasing, so extra distance can be achieved with less effort, but velocity damping is applied to reduce this effect to a maximum of a few meters. Prior research has found small amounts of translation gain can be applied without a perceivable visual effect, but can be used to cover larger distances more easily [23, 25, 26]. A translation gain of 1.5 was applied to this technique for this work. For both the hand-tracked and controller systems, the input for "grabbing" the air to activate the grab-pull mechanic is the same input as grabbing objects in the world to pick them up. If it is not clear to the user if their action will result in motion, simulator sickness may occur [22].

On the Quest 2, hand tracking is performed by machine vision algorithms that process data from four infrared cameras mounted to the headset that are also used to track the controllers. Significant jitter can be observed at all times, with degraded accuracy as the hands become occluded and cut off in the camera view or as they move quickly and become blurred. As a result, several mitigations were developed to reduce the effect of poor hand tracking quality on the locomotion experience. A double-exponential smoothing filter was applied to the raw pinch position at all times. In addition, extra smoothing was applied inversely proportionately to the distance to the headset once the hands were within a certain radius. A common action when using grab-pull to move forwards results in the user's release point ending close to their face. The accuracy at the release point is the most critical, since it also determines the distance traveled due to momentum. By increasing the smoothing at this point, release velocity became more consistent for this type of motion.

3.3.2 *Teleportation and Snap-Turning.*

Other tasks require the user to cover larger distances quickly, such as forming small groups for conversations or playing a larger-scale game. Using grab-pull can only translate by a few meters for each action, so a faster method is useful. In addition, users seated in non-rotating chairs must have a way to virtually turn their avatars. Both of these can be supported by a mechanic that transforms the tracking volume to achieve the desired offset or rotation [3].

In the application, teleportation is achieved by first activating the teleporter. When active, a parabolic trajectory is visualized emerging from an origin and ending at a target position somewhere in the world. If the target position is valid, deactivating the teleporter instantly transports the user to that location. Snap turning is a single activation, left or right, and results in a corresponding 30 degrees rotation of the tracking volume about one of two positions. When the grab-pull system is active, the center of rotation is the hand that is grabbing; otherwise, the play space rotates around the head position.

When using controllers, teleportation is activated by the thumbstick, as is commonly done in consumer games. Holding forward on the thumbstick activates the teleporter arc, which emerges from the controller position. Releasing the thumbstick activates the teleport. Snap rotation activates when pushing the thumbstick left or right, which must be repeated for multiple rotations.

For hand-based teleportation, the same arc-based teleporting ray was used, with the only difference lying in the activation and pointing of the ray. To activate a teleport, the pinch action was reused, as it was found to be the most reliable gesture. However, as it was already in use for grab-based sliding, we needed to multiplex the two actions. Our solution was to reserve three portions of the reachable space for activating the teleporter and snap turns. These regions could not be placed near the center of the viewport since they would interfere with other tasks in the virtual environment, such as picking up objects. The bottom of the viewport is also an inconvenient area to put regions because the hands are often near the bottom of the screen when resting. Therefore, the teleport region was placed forward and above the user's head position, and the snap-turn targets were placed to the sides. Pinching while the user's hands are inside a sphere of a 10 cm radius activates the teleporter, and the user can then lower their hands to point the ray at the desired location on the floor. The direction of the ray is determined by the vector between a point near their head and their hand. If the exact head position were used in this calculation, the ray would be difficult to see since the user is looking directly down the length of the ray. Instead of the head position, a point to the side (depending on which hand activates the teleporter) and below their head was used in this calculation to reduce fatigue by moving the typical release point to a more natural hand position. The choice to use the head-hand vector for the teleporter ray direction differs from the implementation for controllers, which uses the controller direction. This difference was necessary because the accuracy of the orientation of the controllers is far higher and more temporally stable than that of tracked hands due to the presence of the inertial measurement unit (IMU) in the controllers, while the hands rely on purely visual tracking. Additionally, the goal of this study was not to compare hand tracking precision with controller tracking but rather to investigate a realistic implementation of a similar locomotion system for both interaction mechanisms. We believe the chosen inputs represent a high-quality locomotion mechanic that does not interfere with the ability to interact with the world for complex tasks.

4 STUDY DESIGN

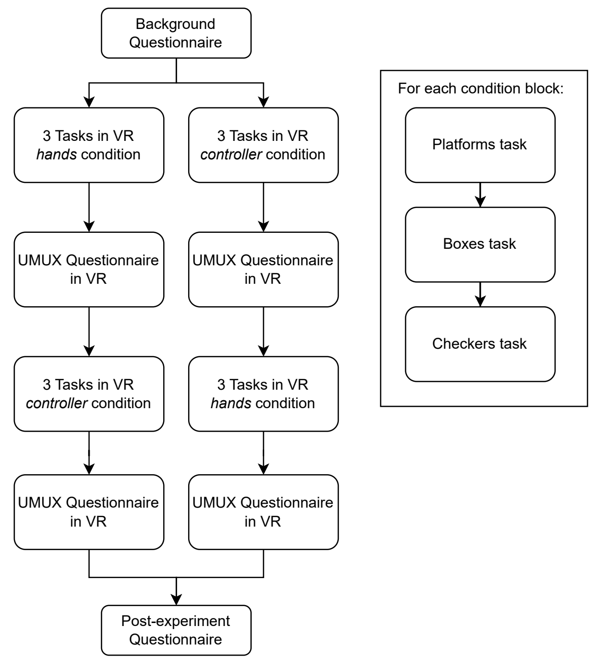

Figure 2: Study procedure. The two participant groups varied in order of the *hands* and *controllers* conditions

We conducted a user study to evaluate the proposed locomotion technique across both hand-tracked (*hands* condition) and controller-tracked (*controllers*) conditions as well as through a variety of tasks. A within-subjects, crossover design was used, such that each participant used both the *hands* and *controllers* condition for each of three tasks as seen in Figure 2. The input condition was alternated between participants, but the three tasks were always performed in the same order first for one condition, then for the other. This was done to combat the learning effect that is common in VR user studies. Performance comparisons between the tasks were not intended, so no order variation was needed. At the beginning of each input condition, participants were instructed how to use each of the locomotion inputs, as well as grab and place a single object. This tutorial procedure took approximately 30 seconds, after which they proceeded to the first task. In general the goal of developing the study environment was not to build a bespoke apparatus for this study, but rather to build a world with various tasks that the user was guided through by the experimenter, who was also in VR with the participant in the same virtual space. For example, there were no artificial limits on where the participant could teleport in the platforms task, and the grab-pull locomotion system was not disabled. Participants were just told that their task was to teleport to each platform in turn, just as would happen in the real world.

4.1 Tasks

Three tasks were designed to test the system in a variety of situations. Each of the tasks represents a different segment of possible application usage.

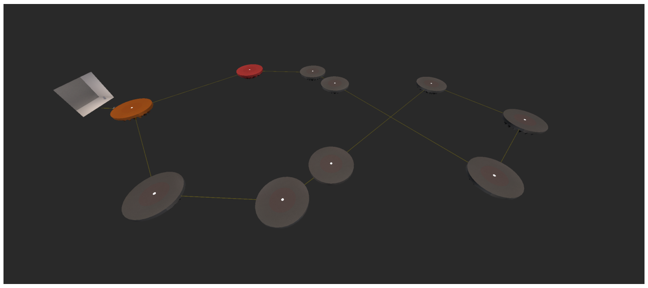

Figure 3: An overview of the platforms task. The layout contains varying distances and turn angles.

4.1.1 Platforms. The first task was intended to focus on measuring the precision and speed of the teleporting portion of the locomotion system. The task consisted of 10 platforms of 2 m radius (Figure 4). All of the platforms were at the same and arranged in such a way that there were a variety of left and right turns of various degrees as well as a variety of distances, from 10 m to 36 m. See Figure 3 for the layout. Due to the parabolic nature of the teleporting ray in both the *hands* and *controllers* conditions, the 36 m distance was near the maximum comfortably possible for the interface. Participants were instructed to perform the course twice in a row. The first time as "precisely as possible" and the second time "as fast as possible," in order to encompass both ends of the speed-precision trade-off. The "fast as possible" trial was always performed second to match the natural increase in speed that happens with growing familiarity with a task. While the grab-pull locomotion technique was not forbidden verbally or through the software, the task was described as teleporting to each platform, and only teleporting data was analyzed.

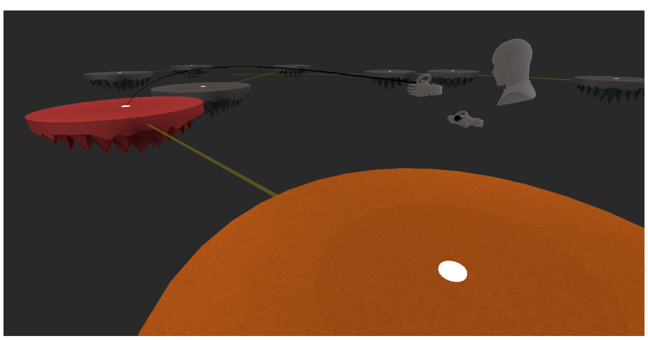

Figure 4: A user teleporting from one platform to the next. The current platform was highlighted in orange, and the next platform was highlighted in red.

4.1.2 Boxes. After completing the platforms task, participants looked to the center of the room, where a set of 10 boxes on posts were spawned (see Figure 5). This task was designed to mimic similar tasks from related works[7, 16] that also attempted to analyze the task performance of various locomotion techniques. One addition beyond that of related work is to randomly set the elevation of the boxes to any value within 1 to 2 m above the ground. There were

Figure 5: A user opening a door of one of the boxes in the boxes task. This is the second of the two box configurations.

two predefined random layouts of boxes, in which the orientation, position, and elevation of the boxes in the room were decided by fixed seeds. The boxes were at least 1 meter apart from each other and were spread in a space of 14x14 meters. To complete the task, participants had to open the door found on one side of the box by grabbing and holding it open while they removed the sphere with the other hand. It was not possible to see whether a sphere had already been removed from the outside, so part of the task was to maintain spatial awareness throughout usage of the locomotion system to travel between boxes. Each box was 60 cm in width, and the spheres inside had a 10 cm radius.

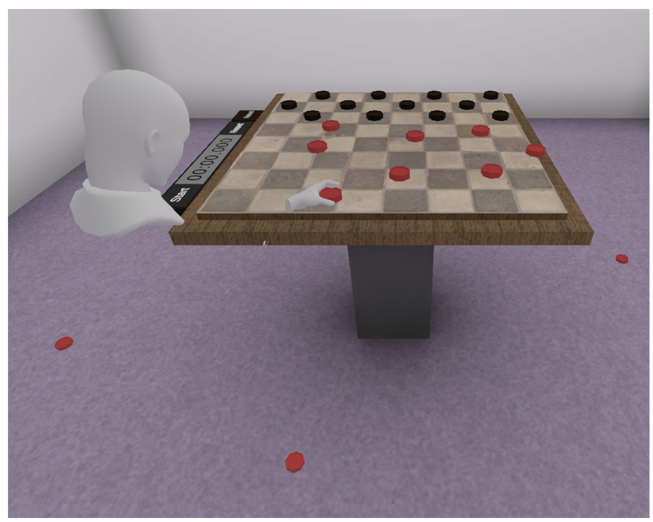

Figure 6: The checkers board in its "messy" state. Several pieces started on the floor.

4.1.3 Checkers. The third task was to clean up a messy checkers board as seen in Figure 6. Participants were presented with a small side room with a checkers board on which the black pieces opposite of them perfectly aligned, but the red pieces were scattered around the board and floor. The task was to pick up all 12 red pieces and place them onto the appropriate squares of the board as demonstrated by the black pieces. There was a physical timer on the side of the board with a start and stop button that the participant started and stopped by themselves. The timer provided motivation to complete the task faster as well as providing another precise UI interaction as part of the task.

4.2 Study Population and Environment

Following human-subjects review approval, 20 participants were recruited from the local university population and lab groups to participate in the study by direct invitation and word-of-mouth. A $15 gift card was used as an incentive, and the entire experience could be completed in 30-45 minutes. Most participants completed the study in our laboratory, but one participant performed the experiment remotely. Communication with this participant happened through a voice call for the portions outside of VR, and used the built-in voice communication in the VR application for communication during the study. No other differences in the run of the study were experienced. All participants were seated in a fixed-rotation chair as in the work by Huber et al. and used a Meta Quest 2 headset [9]. The *controllers* condition was completed with the standard Quest 2 controllers and the *hands* condition with the built-in hand tracking on the Quest 2 using the "MAX" frequency option for hand tracking V2.0 for the best experience on this platform. The experimenter was also in VR to walk the participant through the study, though they always used the controllers. Frame rate inside the application was consistently 90 Hz with no noticeable frame drops.

4.3 Procedure and Measures

After signing the informed consent document, participants filled out a background survey with questions about demographics and prior experience with VR. The necessary buttons or gestures were shown outside of VR for their first condition, then they were instructed to put on the headset and adjust it for a comfortable viewing experience. Once both the participant and the experimenter were inside VR in the same virtual environment, the experimenter demonstrated and explained both grab-pull, teleporting, and snap-turn locomotion techniques. In the *hands* condition, the participant was first asked to calibrate their arm length using the option in the floating tablet interface. After selecting the calibration option, a small box appeared in front of them, which they were told to grab and release at their maximum arm extension. The participant was observed while completing each of the locomotion techniques at least once, and if they were able to successfully complete them, they were instructed to grab a small cube from a table and put it into a box before selecting the separate scene where the actual study would take place. The experimenter guided the participants through each of the tasks, observing their progress just as if the tasks were completed in a physical lab instead of a virtual environment. Every action performed in VR, such as teleport events, grab events, etc., was logged to a file and used for and used to generate the quantitative results for each of the tasks. After completing the third task in VR, a web browser embedded in the environment was used to present survey questions without the context switch that leaving VR would necessitate. These questions followed a standard Usability Metric for User Experience (UMUX) questionnaire [2]. Between the *hands* and *controllers* conditions, the participant was instructed to take off the headset for a few minutes to reduce fatigue from wearing the headset for extended periods of time as well as to provide the same brief introduction to the buttons or gestures for the second condition. After both conditions were completed, another survey was completed at a desktop computer, asking about comparisons between both conditions and their general experience.

5 RESULTS

All participants were able to complete all of the tasks successfully, and no motion sickness, vision problems, or discomfort were reported during the study. The average age of the 20 participants was 22 years (SD=4.2), and 14 reported male, 5 female, and 1 other/prefer not to say. Experience levels were mixed, with 6 participants never having used VR, and 10/20 users reporting 4 or above on a 1-7 scale of VR experience. Comfort playing 3D video games and self-reported competitiveness were similarly mixed.

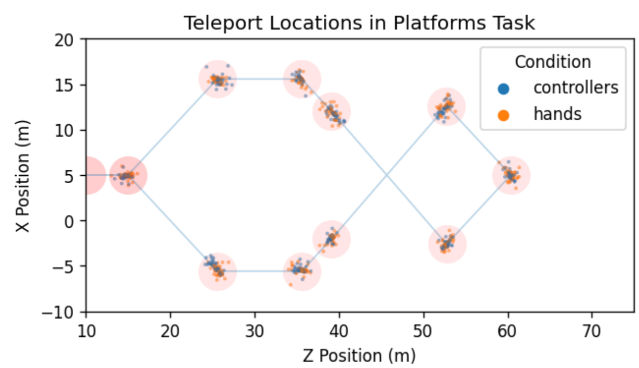

Figure 7: Top-down view of the teleport locations in the *platforms* task. The highest error was in the direction of travel.

Using a Chi-square test for independence, no significant differences were found between the two groups assigned to different condition orders. In addition, none of these background results correlated strongly with the performance results, so they were not used as covariables in subsequent analyses.

A Usability Metric for User Experience (UMUX) questionnaire was asked after both the *hands* and *controllers* conditions. UMUX is a four-question survey designed to generate results comparable to the popular System Usability Survey (SUS) while requiring fewer questions to complete [1, 2]. The *hands* condition generated a SUS-comparable score of 0.62, and the *controllers* condition generated a score of 0.78, which were significantly different according to an independent samples two-tailed t-test (t(18)=-2.39, p=0.02). No difference was found when all of the first trials were compared to the second trials (M1=0.71, M2=0.68, t(18)=0.59, p=0.56).

After both conditions were completed, participants answered four 7-point scale comparison questions between the two conditions as well as two questions between the two primary locomotion inputs. Results for these questions can be found in Figures 9 and 10. Using a one-sample t-test for variation from the mean centered at 4, each of the answers was found to be significant in favor of the controller condition. A Shapiro-Wilk test for normality found significant deviation from normality (p<0.05) in all questions except for the third general condition preference questions (as seen in

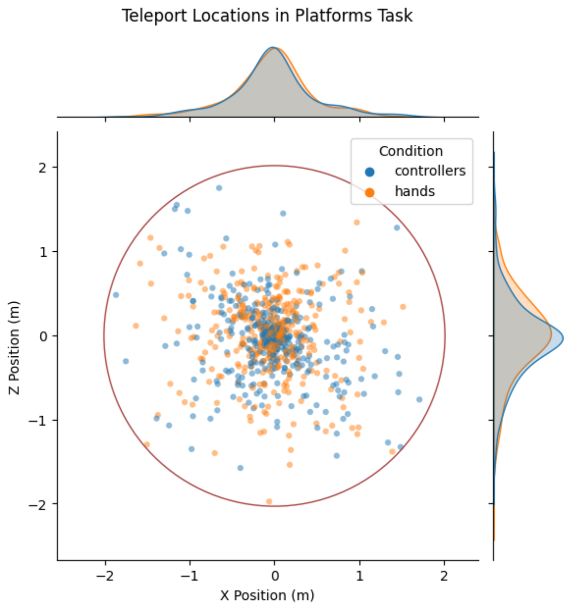

Figure 8: Top down view of all of the teleport locations onto the platforms. Some platforms were approached from different directions, so the direction of highest error is mixed.

Table 1: Summary statistics for all 7-point Likert scale preference questions asked after each condition or after the experiment. 4 was "about the same" and the ends were in favor of either condition. These results can also be seen in Figures 9 and 10

Question	1-samp. t-test
For the Boxes task, which locomotion system did you find more useful?	M=3.85, SD=1.63, t(19)=-0.41, p=0.69
For the Checkers task, which locomotion system did you find more useful?	M=6.25, SD=0.97, t(19)=10.41, p<0.001
Which system did you feel like you could use more quickly/efficiently?	M=5.30, SD=1.38, t(19)=4.21, p<0.001
Which system did you feel was more precise?	M=5.45, SD=1.36, t(19)=4.78, p<0.001
Which system would you prefer to use?	M=4.70, SD=1.42, t(19)=2.21, p=0.04
Which system did you feel took longer to complete?	M=2.30, SD=1.38, t(19)=-5.51, p<0.001

Figure 9 and Table 1), and the locomotion system preference question for the boxes task. Each of these responses with non-normal distributions were also tested with the Wilcoxon signed-rank test centered at a value of 4, and were found not to be symmetric about zero (p<0.05).

Two comparison questions asked about relative preference between the grab-pull and teleport inputs (Figure 10). For the boxes task, no significant preference was found, as participants used a mix of both inputs, but for the checkers task, significant preference was found for the grab-pull input (t(19)=10.41, p<0.001).

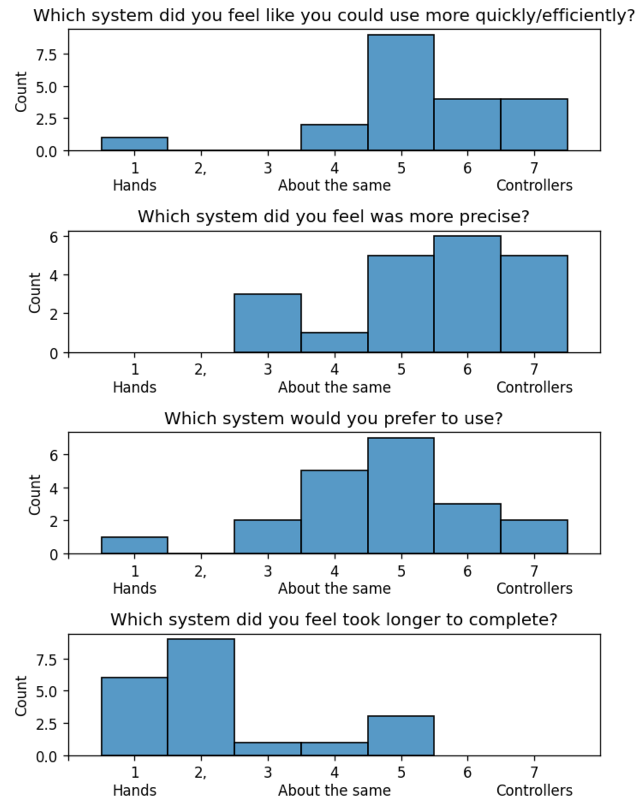

Figure 9: Relative preference scores between the *hands* and *controllers* conditions from the post-experiment questionnaire

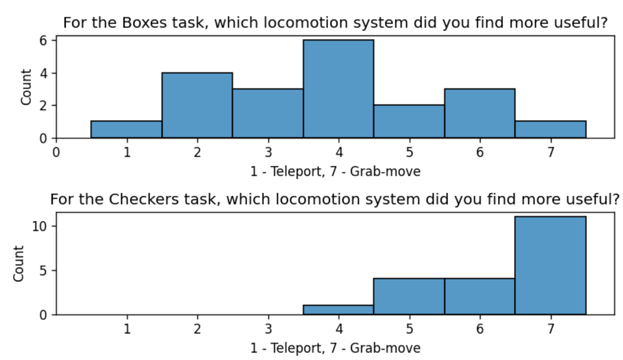

Figure 10: Relative preference scores between Teleport and Grab-move from the post-experiment questionnaire

Quantitative performance was also measured in each of the three tasks. In the platforms task, accuracy was measured as the distance from the center of each platform for each teleport. Overall, No significant difference between the *hands* (M=0.56 m, SD=0.42) and *controllers* (M=0.53 m, SD=0.45) conditions was found (t(18)=0.97, p=0.33), and no improvement was found across the first (M=0.54 m, SD=0.43) and second (M=0.56 m, SD=0.44) trials (t(18)=-0.60, p=0.55).

When isolating the data for only the first run where the participant was instructed to be as precise as possible, significant difference was still not found, but the p-value was reduced to 0.058 (t(18)=1.90) in favor of the *controllers* condition using an independent samples two-tailed t-test. The data for this accuracy measurement can be seen in Figures 7 and 8.

Overall completion time was measured by calculating the time between each of the teleports in the sequence. The *controllers* condition (M=2.93 s, SD=2.82 s) was found to be significantly faster than the *hands* condition (M=6.06 s, SD=5.24 s) using an independent samples two-tailed t-test (t(18)=9.32, p<0.001), and this trend remained when isolating the data to the "fast" run only. No differences (t(18)=-0.81, p=0.42) were found in completion time between the first (M=2.78 s, SD=2.36 s) condition and the second (M=3.00 s, SD=2.64 s) condition, showing no significant learning effect. In order to perform a teleport, a ray must first be activated, then aimed and released. The time with this ray active was measured for both conditions, and again the *controller* condition (M=1.42 s, SD=1.06 s) was found to be significantly faster than the *hands* condition (M=2.84 s, SD=2.27 s, t(18)=10.78, p<0.001) both overall and for the "fast" run only. When limiting "fast" run only, these times reduced to 0.8 s (SD=0.5) and 1.6 s (SD=0.7), but the ratio remained the same. These results can be found in Figure 11.

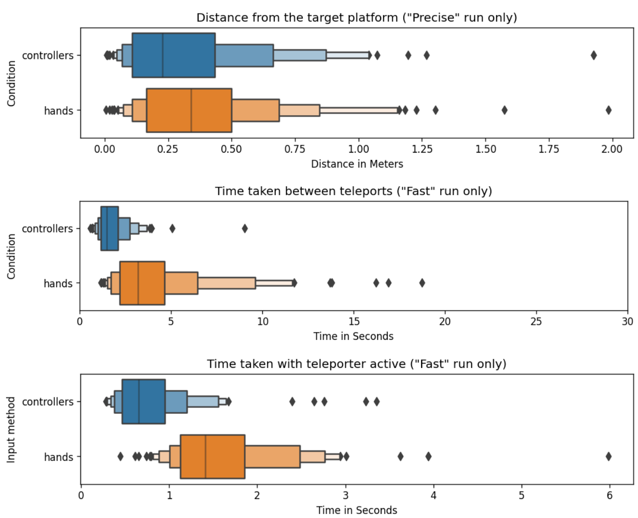

Figure 11: From top to bottom: Distance from the center of the platform for each teleport; Total time between each teleport event. This can be multiplied by 10 for the total task completion time; Time with the teleporting interface active.

Completion time for the boxes task was calculated similarly to the platforms task, with the average time between each box as the primary metric. The *controllers* condition (M=12.42 s, SD=7.75 s) was found to be significantly faster than the *hands* condition (M=16.53 s, SD=8.48 s) using an independent samples two-tailed t-test (t(18)=-4.52, p<0.001) (Figure 12), and and no difference was found between the first condition and the second condition (t(18)=1.43, p=0.15).

Completion time for the checkers task was measured by the physical timer operated by the participants. No significant difference between either the *hands* (M=96.37 s, SD=43.18 s) and *controllers* (M=75.92 s, SD=30.89 s) conditions (t(18)=-1.60, p=0.12) (Figure 13) or the first (M=86.31 s, SD=33.82 s) and second (M=84.85 s, SD=42.80 s) trials were found (t(18)=0.11, p=0.91). Overall average completion time was 86 seconds (SD=39). As each item was picked up and placed in the correct location, the time spent holding each checkers piece was also recorded. Participants held the pieces 0.6 seconds longer on average in the *hands* condition (t(18)=-3.38, p<0.001), but no difference was found between the first and second conditions overall.

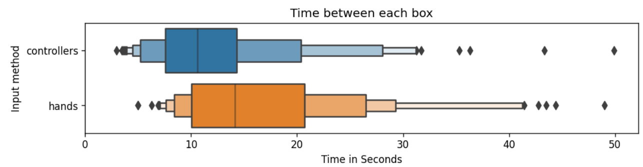

Figure 12: Completion time of the boxes task. The time between each box can be multiplied by the number of boxes for total task time.

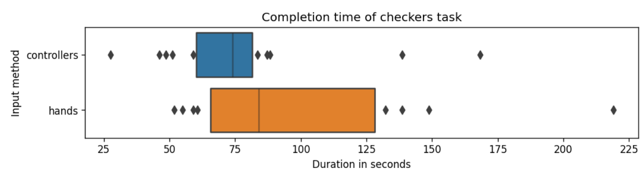

Figure 13: The total completion time for the checkers task. This time is the result of the in-application timer operated by the participant.

On average, participants completed 36.4 (SD=19.4) grab-pull actions and 9.4 (SD=4.5) teleport actions during the boxes task. These averages changed to 43.8 (SD=17.2) grab-pull actions and 2.6 (SD=2.7) teleports for the checkers task.

5.1 User Feedback

Free response feedback was requested for several questions in the post-experiment survey to capture issues or comments not otherwise quantified in the quantitative data or other survey questions. The first such question asked participants to explain why they preferred the teleporting or grab-pull inputs for both the boxes and checkers tasks. 15 out of the 20 participants wrote that the teleport was more useful for long distance movements and the grab-pull input was useful for smaller movements or adjustments. One participant commented, "teleporting was very useful in covering a lot of distance very quickly, but I felt disoriented trying to do it quickly and turning after teleporting was always needed. The grab-pull felt better for moving short distances and it felt more natural physically moving in that way." Only three participants commented on the ability for the grab-pull locomotion to move vertically, whereas the teleporting system could only cover horizontal distance. Five people commented that grab-pull was more natural than teleporting, and six mentioned that grab-pull provided more precise control than teleporting. Another user commented, "Grab and move felt very natural and allowed me to maneuver precisely in the environment." Only one person mentioned fatigue when comparing the two input schemes. None of the responses indicated that the participants perceived the translation gain present in the grab-pull scheme.

Another free-response question prompted participants to discuss the aspect of the tasks or system they found the most difficulty with. The most common issue (seven participants) was the low quality of the hand tracking itself, which resulted in issues performing reliable grab or teleport operations. Five participants found the *hands* condition unreliable at times, especially for the grab-pull technique. Two users mentioned that the high position of the teleport and snap-turn targets in the *hands* condition caused arm fatigue in the platforms task, and two participants commented on the imprecision of the teleporting release, "as I released my fingers to teleport, the teleportation destination shifted whereas with the controllers I did not have this problem."

6 DISCUSSION

Overall, the results showed significant preference and completion time advantages for the *controllers* condition over the *hands* condition, which supports prior work in this area, such as that from Schäfer et al., Masurovsky et al., and Hameed et al. [8, 14, 20]. This research focused more on quantifying that performance delta, and we found 20 to 30% faster completion times for the *controllers* condition.

The three tasks were designed to span the range of applicability of the grab-pull and teleporting interfaces, and we saw this affinity reflected in the results. The platforms task was designed to be the best for teleporting, but the task specified the use of teleporting for movement, so the relative usage in that task is not useful. However participants were able to choose which locomotion system they preferred to use in the boxes and checkers tasks. The grab-pull interface had far more invocations than teleports in both tasks, but the teleporting action results in a longer distance traveled. In the box task, a user could choose to teleport from one of the boxes to another, or use several smaller grab-pull actions to cover the same distance. Since the average number of teleport actions was 9.4 for this task, and only 11 out of the 20 participants used 9 or more teleports (the minimum number of teleports required to travel between 10 boxes), the data supports that many of the transitions between boxes were completed purely using the grab-pull technique. This is corroborated with the qualitative observations of the experimenter, who was watching the participant complete the tasks. Part of the task was remembering which of the boxes had already been visited, as it was not possible to see from the outside whether a box had been completed or not. It was relatively uncommon for a participant to forget which boxes had been visited, but several participants did spend some time moving between boxes to find the one they had not yet completed. Since this only happened a small number of times, it was not possible to quantify which of the input methods resulted in higher occurrences of disorientation.

The average number of teleports in the checkers task decreased compared to the boxes task, while the number of grab-pull actions increased. This supports our hypothesis that grab-pull is a more useful locomotion system for manipulation-heavy tasks that do not require a large range of motion. Preference scores and free-response comments from participants also heavily supported this trend. One behavior that was observed during the study was the use of teleport as a height-reset mechanic. In the checkers task, four of the pieces were on the floor, so participants used the grab-pull technique to move themselves down to the floor to avoid reaching down physically, which would have been especially difficult in a seated position. This advantage in the ability to move more easily in three dimensions is supported in prior work from Lim et al. [12]. When returning to the checkers board to place the piece in the correct location, some participants teleported instead of returning using the same grab-pull technique. The teleport scheme always results in a final vertical position where the physical floor is at the same height as the virtual floor, which is an appropriate height to stand in front of the checkers board to place the piece.

Another trend observed in both the boxes and checkers task was that even in situations where the travel distance was far and the teleporter was used, the grab-pull system was still used to adjust the final position. This would have likely been less common for simpler tasks that required less precision to complete, but the boxes task in this study was specifically designed to require both large-scale movements as well as two-hand interactions at the scale of several centimeters. This was especially noticeable when participants moved to a box, tried to open the door and remove the sphere, but were not able to, then adjusted their position and orientation again using the locomotion system, after which they could more easily complete the task. Qualitative participant responses aligned with this observed behavior.

Another significant difference we found between the *controllers* and *hands* conditions was the increase in usage of the snap-turn mechanic for the *controllers* condition. This difference can likely be attributed almost entirely to the increase in distance the hands need to move to activate the feature. Using controllers, snap turn is activated by pushing the thumbstick to the side, which is a total movement of only a few centimeters by the thumb. For the *hands* condition, the snap-turn activation target is generally tens of centimeters from the current hand location. The interesting result from this data is how much participants chose to not use snap turn as a result, possibly indicating that most of the snap turns in the *controller* condition were not essential to completing the task, though it is still possible the use of snap turn did contribute to a better user experience. Application developers have a large amount of control over how much this feature is used based on the friction that is required to activate it. Snap turns can be avoided entirely by allowing fully physical rotation, but in this study the participant was seated to avoid this.

7 CONCLUSIONS, LIMITATIONS, AND FUTURE WORK

In this work we present a mixed-method locomotion scheme that uses teleportation, snap-turn, and grab-pull locomotion to achieve high task performance in a variety of workloads. We conducted a user study in which participants used this locomotion scheme for three different tasks for both hand-tracked and tracked-controller environments. We analyzed the usage patterns across each of the tasks and across both input conditions, and found significant overall preference for the *controllers* condition as well as faster completion times, which aligns with prior work in this space, though accuracy was less affected. More importantly, we quantified the performance delta that switching to a hand-tracked system incurs.

Though we developed a user study that looked at three tasks spanning a variety of usage patterns, not all use cases are covered by this test. It is possible that some scenarios would result in a larger shift in performance deltas for the *hands* and *controllers* conditions that is not anticipated. Future work could expand this evaluation of this locomotion scheme to further tasks. Another limitation of this study is the relatively small size of the study population and the gender disparity. Future VR applications will be used by all members of society, not just the demographics currently more likely to become involved in a VR research study [18].

The overall performance and quality of hand tracking locomotion systems is highly dependent on the quality of the tracking itself. Our system used Hand Tracking V2.0 on the Meta Quest 2, which is only one of the available tracking systems, which may limit the generalizability of our results. Future headsets or software revisions may also have higher quality hand tracking implementations, changing the results of the study by reducing tracking errors. Alternatively, finger and wrist-worn input devices may offer a compromise between hand-tracking and controller-tracking and should be explored as a means to mitigate hand-tracking and pinch activation issues while maintaining convenience.

ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation under Grant No. 1934481.

REFERENCES

[1] Simone Borsci, Stefano Federici, Silvia Bacci, Michela Gnaldi, and Francesco Bartolucci. 2015. Assessing user satisfaction in the era of user experience: Comparison of the SUS, UMUX, and UMUX-LITE as a function of product experience. *International journal of human-computer interaction* 31, 8 (2015), 484–495.

[2] John J Bosley. 2013. Creating a short usability metric for user experience (UMUX) scale. *Interacting with Computers* 25, 4 (2013), 317–319.

[3] Evren Bozgeyikli, Andrew Raij, Srinivas Katkoori, and Rajiv Dubey. 2016. Point & teleport locomotion technique for virtual reality. In *Proceedings of the 2016 annual symposium on computer-human interaction in play*. 205–216.

[4] Davide Calandra, Michele Billi, Fabrizio Lamberti, Andrea Sanna, and Romano Borchiellini. 2018. Arm Swinging vs Treadmill: A Comparison Between Two Techniques for Locomotion in Virtual Reality.. In *Eurographics (Short Papers)*. 53–56.

[5] Nicola Capece, Monica Gruosso, Ugo Erra, Rosario Catena, and Gilda Manfredi. 2021. A Preliminary Investigation on a Multimodal Controller and Freehand Based Interaction in Virtual Reality. In *Lecture Notes in Computer Science*. Springer International Publishing, 53–65. https://doi.org/10.1007/978-3-030-87595-4_5

[6] Noah Coomer, Sadler Bullard, William Clinton, and Betsy Williams-Sanders. 2018. Evaluating the effects of four VR locomotion methods: joystick, arm-cycling, point-tugging, and teleporting. In *Proceedings of the 15th ACM symposium on applied perception*. 1–8.

[7] Loren Puchalla Fiore, Ella Coben, Samantha Merritt, Peng Liu, and Victoria Interrante. 2013. Towards Enabling More Effective Locomotion in VR Using a Wheelchair-based Motion Platform.. In *EGVE/EuroVR*. 83–90.

[8] Asim Hameed, Andrew Perkis, and Sebastian Möller. 2021. Evaluating Hand-tracking Interaction for Performing Motor-tasks in VR Learning Environments. In *2021 13th International Conference on Quality of Multimedia Experience (QoMEX)*. 219–224. https://doi.org/10.1109/QoMEX51781.2021.9465407

[9] Marlene Huber, Simon Kloiber, Hannes Kaufmann, and Katharina Krösl. 2023. Exploring Locomotion Techniques for Seated Virtual Reality. In *2023 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. IEEE, 925–926.

[10] Chaowanan Khundam, Varunyu Vorachart, Patibut Preeyawongsakul, Witthaya Hosap, and Frédéric Noël. 2021. A comparative study of interaction time and usability of using controllers and hand tracking in virtual reality training. In *Informatics*, Vol. 8. MDPI, 60.

[11] Jinwook Kim, Hyunyoung Jang, Dooyoung Kim, and Jeongmi Lee. 2022. Exploration of the Virtual Reality Teleportation Methods Using Hand-Tracking, Eye-Tracking, and EEG. *International Journal of Human–Computer Interaction* (aug 2022), 1–14. https://doi.org/10.1080/10447318.2022.2109248

[12] Donghae Lim, Shizuka Shirai, Jason Orlosky, Photchara Ratsamee, Yuki Uranishi, and Haruo Takemura. 2022. Evaluation of User Interfaces for Three-Dimensional Locomotion in Virtual Reality. In *Proceedings of the 2022 ACM Symposium on Spatial User Interaction*. 1–9.

[13] Daniel P. Mapes and J. Michael Moshell. 1995. A Two-Handed Interface for Object Manipulation in Virtual Environments. *Presence: Teleoperators and Virtual Environments* 4, 4 (11 1995), 403–416. https://doi.org/10.1162/pres.1995.4.4.403 arXiv:https://direct.mit.edu/pvar/article-pdf/4/4/403/1622823/pres.1995.4.4.403.pdf

[14] Alexander Masurovsky, Paul Chojecki, Detlef Runde, Mustafa Lafci, David Przewozny, and Michael Gaebler. 2020. Controller-Free Hand Tracking for Grab-and-Place Tasks in Immersive Virtual Reality: Design Elements and Their Empirical Study. *Multimodal Technologies and Interaction* 4, 4 (2020). https://doi.org/10.3390/mti4040091

[15] Sophia Neamoniti and Vlasios Kasapakis. 2022. Hand Tracking vs Motion Controllers: The effects on Immersive Virtual Reality Game Experience. In *2022 IEEE International Symposium on Multimedia (ISM)*. IEEE, 206–207.

[16] Thinh Nguyen-Vo, Bernhard E. Riecke, Wolfgang Stuerzlinger, Duc-Minh Pham, and Ernst Kruijff. 2021. NaviBoard and NaviChair: Limited Translation Combined with Full Rotation for Efficient Virtual Locomotion. *IEEE Transactions on Visualization and Computer Graphics* 27, 1 (2021), 165–177. https://doi.org/10.1109/TVCG.2019.2935730

[17] Yun Suen Pai and Kai Kunze. 2017. Armswing: Using arm swings for accessible and immersive navigation in ar/vr spaces. In *Proceedings of the 16th international conference on mobile and ubiquitous multimedia*. 189–198.

[18] Tabitha C Peck, Laura E Sockol, and Sarah M Hancock. 2020. Mind the gap: The underrepresentation of female participants and authors in virtual reality research. *IEEE transactions on visualization and computer graphics* 26, 5 (2020), 1945–1954.

[19] Jussi Rantala, Jari Kangas, Olli Koskinen, Tomi Nukarinen, and Roope Raisamo. 2021. Comparison of Controller-Based Locomotion Techniques for Visual Observation in Virtual Reality. *Multimodal Technologies and Interaction* 5, 7 (2021). https://doi.org/10.3390/mti5070031

[20] Alexander Schafer, Gerd Reis, and Didier Stricker. 2022. Comparing Controller with the Hand Gestures Pinch and Grab for Picking Up and Placing Virtual Objects. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. IEEE. https://doi.org/10.1109/vrw55335.2022.00220

[21] Alexander Schäfer, Gerd Reis, and Didier Stricker. 2021. Controlling Teleportation-Based Locomotion in Virtual Reality with Hand Gestures: A Comparative Evaluation of Two-Handed and One-Handed Techniques. *Electronics* 10, 6 (mar 2021), 715. https://doi.org/10.3390/electronics10060715

[22] Craig R Sherman. 2002. Motion sickness: review of causes and preventive strategies. *Journal of travel medicine* 9, 5 (2002), 251–256.

[23] Betsy Williams, Gayathri Narasimham, Tim P McNamara, Thomas H Carr, John J Rieser, and Bobby Bodenheimer. 2006. Updating orientation in large virtual environments using scaled translational gain. In *Proceedings of the 3rd Symposium on Applied Perception in Graphics and Visualization*. 21–28.

[24] Preston Tunnell Wilson, William Kalescky, Ansel MacLaughlin, and Betsy Williams. 2016. VR locomotion: walking> walking in place> arm swinging. In *Proceedings of the 15th ACM SIGGRAPH Conference on Virtual-Reality Continuum and Its Applications in Industry-Volume 1*. 243–249.

[25] Xianshi Xie, Qiufeng Lin, Haojie Wu, Gayathri Narasimham, Timothy P McNamara, John Rieser, and Bobby Bodenheimer. 2010. A system for exploring large virtual environments that combines scaled translational gain and interventions. In *Proceedings of the 7th Symposium on Applied Perception in Graphics and Visualization*. 65–72.

[26] Ruimin Zhang, Bochao Li, and Scott A Kuhl. 2014. Human sensitivity to dynamic translational gains in head-mounted displays. In *Proceedings of the 2nd ACM symposium on Spatial user interaction*. 62–65.

[27] Jingbo Zhao, Ruize An, Ruolin Xu, and Banghao Lin. 2022. Comparing hand gestures and a gamepad interface for locomotion in virtual environments. *International Journal of Human-Computer Studies* 166 (oct 2022), 102868. https://doi.org/10.1016/j.ijhcs.2022.102868

[28] Daniel Zielasko, Sven Horn, Sebastian Freitag, Benjamin Weyers, and Torsten W Kuhlen. 2016. Evaluation of hands-free HMD-based navigation techniques for immersive data analysis. In *2016 IEEE Symposium on 3D User Interfaces (3DUI)*. IEEE, 113–119.