Article

Comparing and Contrasting Near-Field, Object Space, and a Novel Hybrid Interaction Technique for Distant Object Manipulation in VR

Wei-An Hsieh [1], Hsin-Yi Chien [1], David Brickler [2], Sabarish V. Babu [2] and Jung-Hong Chuang [1,*]

[1] Department of Computer Science, National Yang Ming Chiao Tung University, Hsinchu 30010, Taiwan; b04902106@ntu.edu.tw (W.-A.H.); hsinyichien@icloud.com (H.-Y.C.)
[2] School of Computing, Clemson University, Clemson, S.C. 29634, USA; dbrickl@g.clemson.edu (D.B.); sbabu@clemson.edu (S.V.B.)
[*] Correspondence: jhchuang@cs.nycu.edu.tw; Tel.: +886-3-5731829

Abstract: In this contribution, we propose a hybrid interaction technique that integrates near-field and object-space interaction techniques for manipulating objects at a distance in virtual reality (VR). The objective of the hybrid interaction technique was to seamlessly leverage the strengths of both the near-field and object-space manipulation techniques. We employed bimanual near-field metaphor with scaled replica (BMSR) as our near-field interaction technique, which enabled us to perform multilevel degrees-of-freedom (DoF) separation transformations, such as 1~3 DoF translation, 1~3 DoF uniform and anchored scaling, 1 DoF and 3 DoF rotation, and 6 DoF simultaneous translation and rotation, with enhanced depth perception and fine motor control provided by near-field manipulation techniques. The object-space interaction technique we utilized was the classic Scaled HOMER, which is known to be effective and appropriate for coarse transformations in distant object manipulation. In a repeated measures within-subjects evaluation, we empirically evaluated the three interaction techniques for their accuracy, efficiency, and economy of movement in pick-and-place, docking, and tunneling tasks in VR. Our findings revealed that the near-field BMSR technique outperformed the object space Scaled HOMER technique in terms of accuracy and economy of movement, but the participants performed more slowly overall with BMSR. Additionally, our results revealed that the participants preferred to use the hybrid interaction technique, as it allowed them to switch and transition seamlessly between the constituent BMSR and Scaled HOMER interaction techniques, depending on the level of accuracy, precision and efficiency required.

Keywords: near-field interaction technique; far-field interaction technique; hybrid interaction technique; distant object manipulation; empirical evaluation

1. Introduction

The availability of cost-effective, user-friendly, and powerful hardware in virtual reality (VR) has led to a surge in applications, such as in gaming, training, engineering, design, social activities, and education. To interact with a virtual environment (VE), it is essential to be able to manipulate virtual objects within that environment. For a long time, researchers have been exploring ways to manipulate objects in virtual settings, but this is still a challenging endeavor. Despite considerable research, there is still a need to build a natural virtual interface with the accuracy and effectiveness required for professional purposes, such as in product design. The current techniques do not provide 9DoF transformations with DoF separation in manipulations, which could potentially enhance accuracy and enable more flexible operation for complex tasks.

One way of manipulating an object in VE is to select and manipulate the object from a distance. This indirect manipulation allows the user to pick up an object outside their arm's reach and interact with it, without having to move towards it within the VE.

However, manipulating objects in this way may magnify the hand manipulation error due to human instability, because with these distant manipulation techniques, the movement of the object scales up the farther away it is, increasing the error of object placement with distance [1]. Velocity-based scaling has been proposed to reduce this scale-up error [1–3]. The next approach to improving accuracy is DOF separation [4,5]. It has been stated that complete DoF separation (1DoF translation and 1DoF rotation) through virtual widgets can prevent unwanted transformations and improve the accuracy, precision, and granularity of placement, at the cost of increasing the time needed for complex tasks [4,5].

A typical technique for manipulating distant objects is Scaled HOMER [1], which leverages HOMER [6] with PRISM [3], with the objective of improving the accuracy of HOMER using velocity-based scaling. Although the scale-up error can be reduced using velocity scaling, Scaled HOMER still suffers from poor vision and imprecise motion control, and hence higher motion instability.Moreover, it only offers 6DoF simultaneous translation and rotation, making it more suitable for coarse transformations.

In this contribution, we propose a hybrid interaction interface that integrates the bimanual near-field metaphor with scaled replica (BMSR) technique that we presented previously in [7] and the popular Scaled HOMER [1] interaction technique. BMSR uses a scaled replica placed within arm's reach to manipulate its far-field counterparts [7]. Manipulation of the replica using its bounding box primitives leads to an intuitive interface and enables implementation of 1~3DoF translation, 1~3DoF uniform and anchored scaling, and 1DoF rotation. These options create an interface that supports 7 degrees of freedom, in general, for precise manipulation. Additionally, 3DoF rotation and 6DoF simultaneous translation and rotation are also supported. Supporting multilevel DoF separation can increase the precision of manipulation [5] and also offer more flexibility for manipulation. A key factor of this bimanual near-field interaction technique is that it is an indirect method, but the scaled replica is directly manipulated at an arm's-reach distance, taking advantage of fine motor movement control, better depth presentation and perception [8], and enhanced vision during personal space interactions. We also conducted a comparative study to understand to what extent the objective performance and subjective impressions and perception of the participants differed between the near-field (BMSR with scaled replica), object-space (Scaled HOMER), and hybrid techniques for distant object manipulation in VR.

Our paper highlights two key contributions. First, we proposed a hybrid interface that aims to balance the accuracy and precision of the BMSR with the rapid long-range movements of the Scaled HOMER technique. Second, we conducted a novel repeated measures comparative evaluation to determine how these interaction techniques—near-field BMSR, far-field or object-space Scaled HOMER, and our hybrid technique—affected users' objective performance variables and subjective impressions of the three interaction techniques for distant object manipulation in VR.

2. Related Work

It is possible to manipulate objects in virtual reality (VR) by directly grabbing and manipulating objects within arm's reach. Examples of these techniques include simple virtual hand [9], Air-TRS [10], spindle [11], handle bar [12,13], Spindle+Wheel [14], crank handle [13], grasping object [13], 6-DoF hand [15], 3-DoF hand [15], widgets [5], and PinNPivot [16]. The user must approach objects that are beyond their reach before manipulating them. Transitioning between manipulating objects and navigating the virtual environment (VE) can be disruptive to the user's experience, even when teleportation is employed as a common locomotion technique [17]. A further problem is that some existing methods only provide 6DoF simultaneous translation and rotation, while others only support a restricted set of DoF separation transformations, such as 3D translation, 3D rotation, and 3D uniform scaling. Recently, PinNPivot has been proposed, offering a more extensive range of transformations, such as 3D translation, 1–3D rotation, and 6DoF simultaneous translation and rotation [16]. None of the current techniques provide full

9DoF manipulation [18] and 1–3D anchored scaling. However, when supported by a high DoF transformation, direct manipulation can mimic interactions in the physical realm [19]. When an object is within arm's reach, users have a clear view of it and a good understanding of its location due to proprioception, giving them a greater sense of control [17]. Spatial relationships between the target object and the objects closest to it can be occluded, making accurate manipulation impossible. To provide more accurate direct manipulation, Choi et al. proposed providing the user with auxiliary views from various viewpoints [20].

The handle box and the tBox widgets have been used to manipulate objects directly on mouse and keyboard interfaces [21,22]. Handle box is a bounding box that encompasses the object, with a lifting handle to move it up and down, and four rotation handles to rotate it around its central axis [21]. The tBox widget consists of a wireframe box surrounding the target object, with which the user can drag an edge to move the object along the axis containing the edge or drag a face to rotate the object [22]. Recently, we released BMSR [7], which features a bounding box widget that allows users to translate or scale distant objects in 1D, 2D, and 3D, by dragging the faces, edges, and vertices of the bounding box, and to rotate in one dimension by grabbing a handlebar and dragging an edge of the box.

A second type of interface enables users to manipulate objects from a distance. This kind of manipulation allows the user to interact with objects that are beyond their arm's reach without having to move around within the virtual environment. In the mid-1990s, two techniques were proposed for this purpose: Go-Go [23] and ray casting [24]. Go-Go utilizes a method that increases the user's arm length and applies nonlinear mapping for interacting and manipulating distant objects, while ray casting involves the user selecting an object with a ray and manipulating the object that is attached to the end of the ray. As shown in [6], Go-Go, stretch Go-Go, and ray casting had considerable drawbacks. Consequently, HOMER (hand-centered object manipulation extending ray casting) was proposed. This technique uses ray casting to select an object and then attaches a virtual hand to it, allowing the user to manipulate the object with the virtual hand. The scaling is based on the distance between the user's body and the hand and the distance between the user's body and the object. This scaling factor can amplify the input and magnify the error in hand manipulation [1]. To reduce the scaled error, the PRISM techniques [2] were designed to reduce object movement, for greater accuracy when the hand moves slowly. Wilkes et al. proposed Scaled HOMER [1], a combination of PRISM and HOMER, to enhance performance in manipulation tasks that require a high degree of precision. This method utilizes velocity-based scaling. Scaled HOMER can increase the accuracy of manipulation by decreasing the scaled error; however, its manipulation error remains the same due to the nature of distal operations. Additionally, it can cause problems of inaccurate depth perception and blurred vision when looking at distant objects. Furthermore, it only supports 6DoF simultaneous translation and rotation, which is often suitable for large or rough transformations [18]. In addition to velocity scaling, an adaptive gain approach was recently proposed to improve the accuracy and efficiency of distant object manipulation [25]. Gains are calculated through fitting user data collected during object manipulation.

One way to improve the precision of manipulation is to separate the degrees of freedom (DoF separation) [4,5]. Mendes et al. compared simple virtual hand (which has 6DoF simultaneous translation and rotation), PRISM (which has 6DoF simultaneous translation and rotation with velocity-based scaling), and a widget for full DoF separation. The use of widgets to achieve complete DoF separation has been observed to result in higher accuracy, although it can take longer to complete complex tasks [5]. DoF separation has been applied to direct manipulation techniques such as widget [4,5] and PinNPivot [16]; however, it has not been used with distal manipulation methods.

The third way of manipulating objects from a distance in virtual reality is to manipulate them indirectly by manipulating scaled replicas in the user's near field. World-in-Miniature (WIM) [26] is a scaled-down representation of the entire environment, offering users a comprehensive overview of the environment, a convenient way to select and manipulate objects, and the ability to teleport. However, its primary purpose is not to provide precise

manipulation. Pierce et al.'s voodoo dolls [27] allow users to manipulate the target object's doll with their dominant hand, while keeping the dolls for the context objects in their non-dominant hand. This takes advantage of a division of labor between the dominant and non-dominant hands [28], in which the dominant hand of the users works within a reference frame set up by their non-dominant hand [28]. This offers a convenient way to interact with objects; however, it may be affected by precision problems, due to the instability of controlling both hands and executing 6DOF simultaneous translation and rotation. The near-field interface with scaled replicas (BMSR) proposed by us in [7] aims to improve the accuracy of manipulating distant objects using two mechanisms. First, it manipulates a scaled replica in the near field, instead of its counterpart in the object space, and hence is able to take advantages of finer motion control and clear vision in near-field manipulations. As a result, the manipulation error is reduced and the manipulation precision is increased. Second, its support for multilevel DoF separation may increase the manipulation precision [5] and offers more flexibility for complex tasks. However, for long-range translation, the near-field interface may require the user to select and move the object multiple times.

The strengths and weaknesses of the current approaches are varied, and none of them are capable of dealing with manipulations that require different levels of precision. Integrating different techniques could potentially take advantage of the benefits of component techniques. However, to the best of our knowledge, only a few hybrid techniques have been proposed. HOMER could be considered a hybrid technique that integrates the ray casting technique and the virtual hand technique [6] to manipulate distant objects. More recently, ReX Go-Go (an enhanced Go-Go) and rabbit-out-of-the-hat WIM (an enhanced WIM) were combined to facilitate precise selection of distant targets in dense and occluded virtual environments [29]. Our research goes a step further by integrating the popular Scaled HOMER distal interface [1] with the BMSR near-field interface [7], with the goal of taking advantage of the benefits of both and thus satisfying different accuracy, precision, and efficiency requirements.

3. Hybrid Interaction Techniques

In this section, we describe the three interaction techniques included in our comparative empirical evaluation. The first is BMSR, a bimanual near-field interface proposed by Lee et al. [7] that manipulates a scaled replica of the selected object in the near-field of the user. The second interface is the well-known Scaled HOMER [1], which manipulates distant objects using an attached virtual hand and utilizes velocity-based scaling to improve accuracy. The third is the proposed hybrid interaction technique that integrates BMSR and Scaled HOMER to allow users to seamlessly alternate between the two as the situation demands. Both BMSR and the Scaled HOMER are used as a baseline for comparison.

With all three interfaces, object selection is handled using ray casting. When the user points to an object using ray casting, the object is highlighted with a contour glow. Once the object has been selected, it is highlighted for confirmation.

3.1. Bimanual Near-Field Interface with Scaled Replica

In the bimanual near-field technique with scaled replica (BMSR) [7], when an object is selected, a scaled replica of that object known as the target replica is created and a bounding box associated with the target replica is constructed. The target replica is then placed in front of the user, within arm's reach. The size of the target replica is scaled to 20 cm, as mentioned in [7]. By having a replica of the target in their close vicinity, the user can gain a better understanding of the object's location through proprioception and improved depth perception, which should lead to a greater sense of control over the object [17].

Based on the principles of division of labor [28] and symmetric or asymmetric movement of two hands, this approach allows supporting rotation, uniform scaling, and anchored scaling with two hands. The interaction using the bounding box allows for a

convenient and intuitive interface and supports unimanual 1D~3D translation, bimanual 1D~3D scaling, and 1D rotation.

The transformations for the object are based on which primitives of the bounding box the user selects. First, by grabbing and moving one of the faces on the box, the user can perform 1D translations along the direction that is perpendicular to that face (see Figure 1a). Second, grabbing one of the edges allows 2D translations along the plane that is perpendicular to the edge (see Figure 1b). Finally, grabbing one of the box's vertices allows for 3D translation (see Figure 1c).
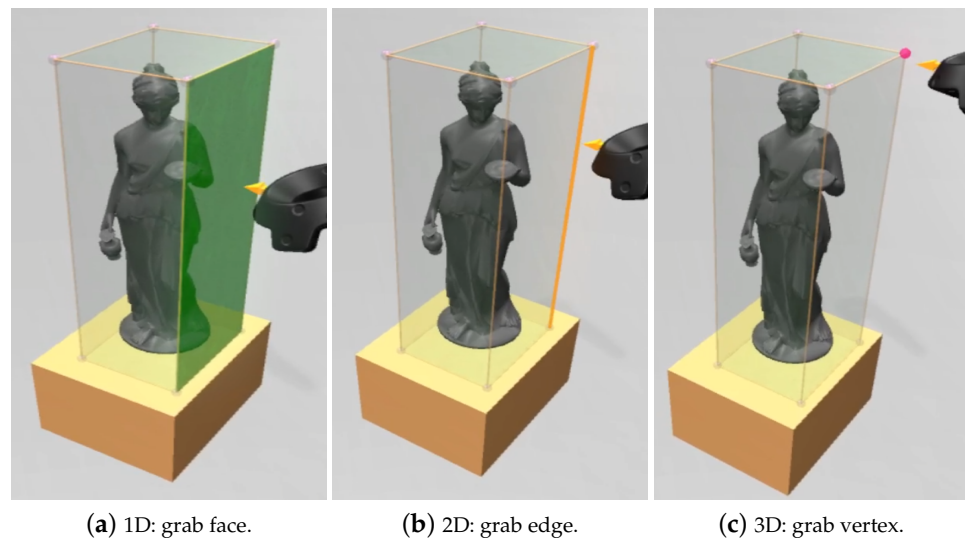

(**a**) 1D: grab face.   (**b**) 2D: grab edge.   (**c**) 3D: grab vertex.

Figure 1. Unimanual technique for translation.

Rotation follows the principle of bimanual division of labor [28]. The user grabs an edge with their dominant hand and then uses their non-dominant hand to grab a handle bar positioned at the center of one of the faces perpendicular to the grabbed edge. The edge or handlebar can be grabbed in any order. When this handle bar is grabbed, the user can then rotate the replica using that handle bar as the rotation axis (see Figure 2a–c).
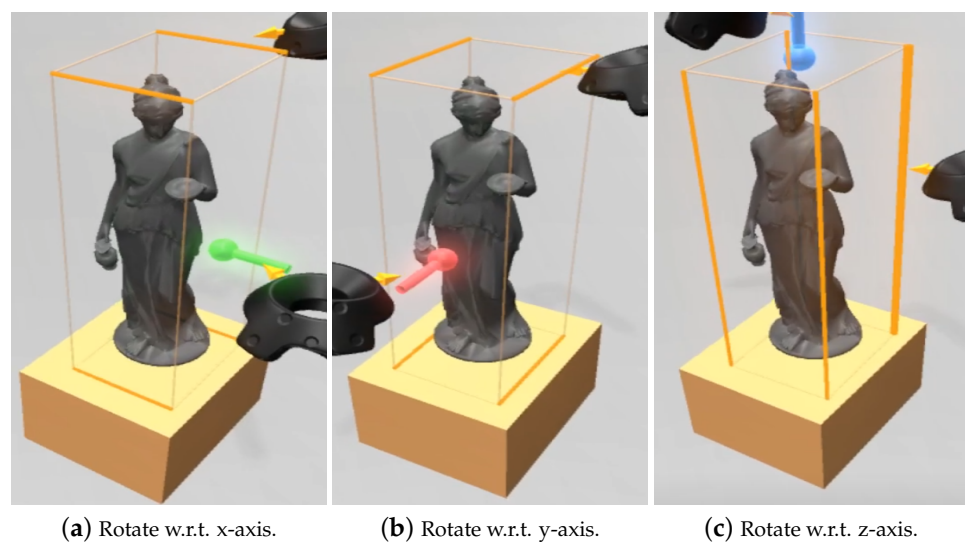

(**a**) Rotate w.r.t. x-axis.   (**b**) Rotate w.r.t. y-axis.   (**c**) Rotate w.r.t. z-axis.

Figure 2. Bimanual technique for 1DoF rotation.

High-DoF transformations provide more natural but less accurate object manipulation [18] and hence can be used for rough transformations. Separation of degrees of

freedom in transformation can provide better precision and prevent unwanted transformations [5]. In practice, it is desirable to perform a rough transformation using high-DoF transformations before performing more accurate transformations using DoF separation operations. In addition to 1D~3D translation, 1D~3D scaling, and 1D rotation, Lee et al. also implemented 6DoF simultaneous translation and rotation. We also implement our own function of 3DoF rotation with respect to the center position. For 6DoF simultaneous translation and rotation, the user directly grabs the scaled replica to move and rotate it freely in virtual space (see Figure 3a). For 3DoF rotation with respect to the center position, we follow a similar interface for 6DoF translation and rotation; that is, the user uses their nondominant hand to grab and hold the scaled replica and uses their dominant hand to directly grab the scaled replica to rotate it (see Figure 3b).
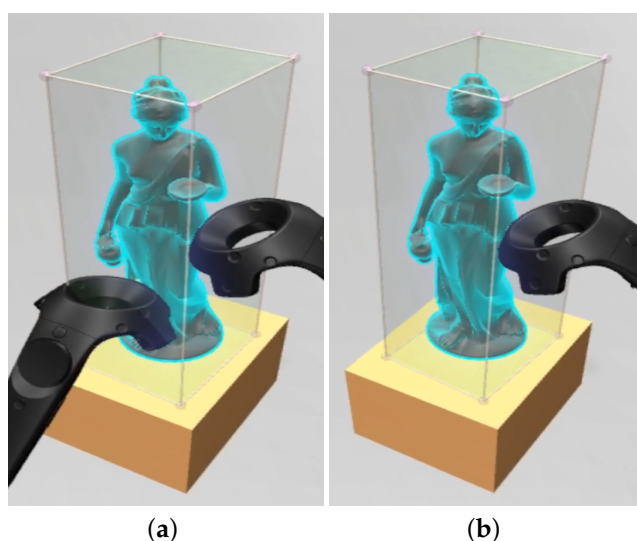

(**a**) (**b**)

Figure 3. 3DoF rotation and 6DoF simultaneous translation and rotation. (**a**) 3DoF rotation; (**b**) 6DoF simultaneous translation and rotation.

As for the uniform scaling and anchor scaling described in [7], we could not use them in our experiment because Scaled HOMER does not include a comparable scaling function.

3.2. Scaled HOMER

Bowman et al. combined Go-Go [23] and ray casting [24] into a distal interaction technique called HOMER (hand-centered object manipulation extending ray casting) [6], which provides better controllability over Go-Go or ray casting alone. Later, HOMER was leveraged with PRISM [3] to form Scaled HOMER [1], with the aim of using velocity-based scaling to improve accuracy and precision. The results showed that Scaled HOMER outperformed HOMER in both precision and efficiency. The reason why we chose Scaled HOMER as our object-space manipulation technique is that Scaled HOMER features both rapid translation and precise control.

3.3. Hybrid Interaction Interface

A *hybrid* approach is proposed to integrate BMSR and Scaled HOMER [1]. This is because both interfaces use ray casting for object selection, making it easier to combine them. Additionally, the two interfaces are complementary to each other. Scaled HOMER's 6DoF simultaneous translation and rotation is well suited for long-distance movements and coarse transformations, allowing for quick and natural interactions [1]. In comparison, BMSR may require the user to select and move the object multiple times to cover the same translation distance. It has been reported that when using Scaled HOMER to perform a basic 1D or 2D translation or rotation around a chosen axis, there may be undesired transformations [5]. Nevertheless, BMSR is capable of dealing with these transformations

in a much more effective manner. The velocity-based scaling in Scaled HOMER is only useful for translation, not rotation, and even then, the improvement in accuracy is limited. Users can have difficulty making small and subtle adjustments, due to the nature of velocity-based scaling and 6D of simultaneous translation and rotation. Furthermore, when objects are far away, it can be difficult to discern how the object is related to its context and how far away it is. We predicted that BMSR would be beneficial for improving the accuracy and precision in both translation and rotation tasks, due to DoF separation, fine motor control, and clarification of vision in the near-field space.

As demonstrated in Figure 4, when aiming at an object with ray casting, the user can press the trigger briefly to select the object and enter BMSR mode. If they instead hold the trigger for more than 0.32 s before releasing it, the simulation will select the object and enter Scaled HOMER mode. When in Scaled HOMER mode, the user can switch to BMSR mode by pressing the trigger briefly. To switch back to Scaled HOMER mode, the user holds the trigger for more than 0.32 s and then releases it. To exit any mode and deselect the object, the user presses the touchpad on the controller.

To determine the time duration of the trigger press for entering Scaled HOMER mode, we asked ten people to press and release a trigger button 100 times, making sure the button was fully released before starting the next round. We measured the time taken from the first press to the last release, and the average time was 0.28 s, with a standard deviation of 0.02 s. We concluded that if someone held the trigger button for more than 0.32 s, they intended to switch to the Scaled HOMER interface in the hybrid interface.

The hybrid interface allows users to benefit from the best of both Scaled HOMER and BMSR, with minimal effort to switch between them. When a coarse transformation or long-distance translation is needed, the Scaled HOMER mode can be used. For more precise manipulation, the BMSR mode is the way to go. This quick and easy transition between the faster Scaled HOMER and the more accurate BMSR can help balance between speed and accuracy [30].
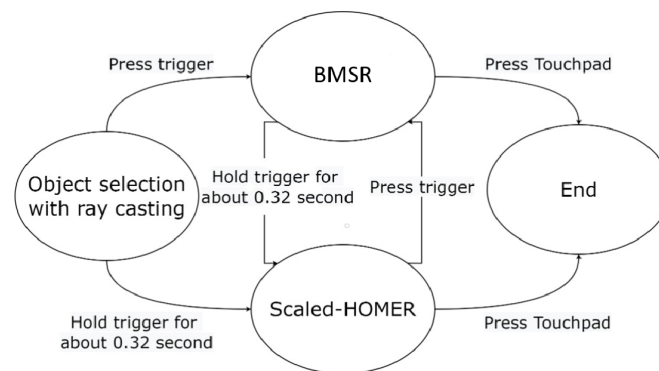

Figure 4. Transition graph for hybrid interaction interface.

4. User Study

We conducted a within-subjects study to compare the BMSR, Scaled HOMER, and proposed hybrid interaction techniques. For fairness and consistency of the comparative study, in the BMSR, we ignored the functionality of uniform scaling and anchored scaling, as Scaled HOMER has no such function to compare with.

Our research question was as follows:

RQ: *To what extent did participants' objective performance, subjective impressions, and perceptions differ between the near-field (BMSR with scaled replica), object-space (Scaled HOMER) and hybrid interactions techniques for object-space or distant-object manipulation in VR?*

To address this research question, we formulated the following hypotheses:

H1: *We hypothesized that the BMSR would outperform Scaled HOMER in accuracy.*
H2: *We hypothesized that Scaled HOMER would outperform BMSR in economy of movement.*
H3: *We hypothesized that Scaled HOMER would result in faster movement times than BMSR.*

H4: *We hypothesized that the hybrid method would outperform Scaled HOMER or BMSR in speed, accuracy, and economy of movement.*

The basis for H1 and H2 was the fact that BMSR allows users to view and manipulate a scaled replica of the object they are moving in their personal space, which allows enhanced depth presentation and perception and enhanced proprioception, which tends to facilitate greater precision and motor control, and potentially reduces hand instability issues [31–33]. On the other hand, Scaled HOMER is a remote manipulation technique for distant object manipulation. Hence, the human instability for manipulation was expected to be larger than that of BMSR. Although Scaled HOMER can be prone to exaggerations in movement error for distant objects, as it utilizes velocity-based scaling, it leverages gross motor movements characterized by short and fast movements, in order to manipulate distant objects [34]. Therefore, with regard to H3, we expected that Scaled HOMER would result in faster movements compared to BMSR. However, BMSR offers multiple levels of degrees-of-freedom (DoF) separation, being capable of leading to more precise object manipulation, while the simultaneous 6DOF translation and rotation used in Scaled HOMER was expected to have an advantage in broader gross or coarse movements in object space manipulations [4,5].

With the hybrid interaction technique, users can switch freely between each of the two modes, potentially leveraging the advantages of both. This method of combining the advantages of different interactions has been shown to be generally effective in improving performance, and specifically for selection and manipulation accuracy [35,36]. Thus, with regards to H4, we expected the hybrid interaction technique to balance between accuracy and efficiency.

4.1. Participants and Apparatus

Using G* Power, we computed an a priori power analysis to determine the number of participants in our study. Using an effect size = 0.25, α = 0.05, power (1-β) = 0.95, number of groups = 3, total number of measurements = 12, and correlation among repeated measures = 0.5, we determined a sample size of 18 participants. Thus, we conducted a user study with 18 participants recruited through a Facebook recruitment page. Using a balanced Latin square design, we assigned participants randomly to one of 3 orders of conditions. Each experimental condition appeared in each of the 3 orders in either the first, second, or third sessions. Therefore, we had a total of 6 participants randomly assigned to each of the 3 orders of the experimental sessions, as per the balanced Latin squares design. Of the participants, 10 were male, 7 female, and one did not disclose their gender. All participants were between the ages of 18 and 40 years and were avid gamers, playing on a PC or smartphone. Of the 18 participants, ten had previous experience with a VR system. The majority of them had used an HTC Vive headset, while one had used Google Cardboard.

For the study, computers with an NVIDIA GeForce 1080 GPU and HTC Vive Pro headsets were used. Participants used the trigger button on the HTC Vive controller to select an object and pressed the controller's touchpad to deselect an object. The experiment was carried out in three sessions conducted over a period of three separate days, to minimize or eliminate the effect of learning or carryover, in a manner similar to [33,37]. Each day, participants were randomly assigned to one of three different interaction techniques (Scaled HOMER, BMSR, and hybrid interface). A Latin square design determined the order of the conditions that the participants experienced.

4.2. Tasks

In our user study, participants were asked to complete three different types of tasks: pick-and-place [32], docking [38–40], and tunneling tasks [41]. These tasks have been well established in the 3D user interface literature for comparative evaluation of interaction techniques for manipulation-type performance. Additionally, similar tasks were also used for the evaluation of near-field and object-space interaction techniques in the IEEE 3D

User Interface Conference 3DUI Contest in 2016 [42]. For each of these tasks, trials were presented as a random permutation of two variables; namely, distance from object to user (medium and far) and object size (medium and large). Therefore, each participant had to perform four tests for each type of task.

There were two steps in the pick-and-place task (as in [32]). At the beginning, there was a semi-cylinder that appeared in the air above a plane, and participants were tasked with placing it into a hole on the plane. In Step 2, participants were tasked with placing the semi-cylinder object in a concave groove object, such that the semi-cylindrical object fit perfectly into the convex groove, which was a situation on a planar surface, as depicted in Figure 5a.

There were also two steps in the docking task (as in [38–40]). First, a pyramid with five different colored spheres in each vertex was initialized on the plane. There was also a reference pyramid-shaped wireframe target that was presented with the same five different colored spheres in each vertex in a different pose in the scene. Participants needed to dock the wireframe target pyramid to match the color of each vertex through a combination of translation and rotation manipulations of the target, as they tried to overlay it onto the reference wireframe pyramid object perfectly. At the beginning of the task, the reference wireframe pyramid object appeared in the plane and the wireframe reference pyramid appeared in the air for Steps 1 and 2, respectively, as shown in Figure 5b.

In the tunnel task (as in [42]), we made three tunnels, each with a predetermined entrance and exit. The first two tunnels were straight, and the last one was C-shaped. All three tunnels were rendered with appropriate color and transparency. Participants had to insert a cube through three tunnels in sequence, as shown in Figure 5c. The size of the tunnels was slightly larger than that of the cube, so the participants had to constantly adjust the translation and orientation of the cube and carefully maneuver it so that it passed through the tunnels, minimizing collisions with the tunnel walls, while moving the cube through the tunnels and completing the task as accurately and quickly as possible.

With these three tasks, we could gain an understanding of how the precision of the BMSR technique could help to reduce the number of collisions and how the rapid movements of the Scaled HOMER could translate the object in large scale through the pick-and-place task. Performing a docking task is highly dependent on the efficiency and accuracy of the technique. The tunnel task has a strong emphasis on guiding the object through the tunnel without colliding with the tunnel walls, which would naturally require a high degree of motor control and precise movements to perform successfully. As such, we hypothesized that the BMSR and hybrid interfaces would yield fewer collisions than Scaled HOMER. There was multi-modal audio and visual feedback when manipulated objects collided with either the plane, target, or tunnel in each task.
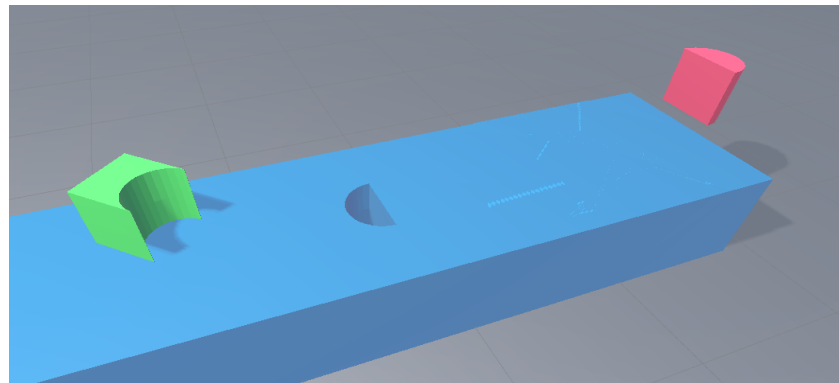
(**a**) Pick and place task.
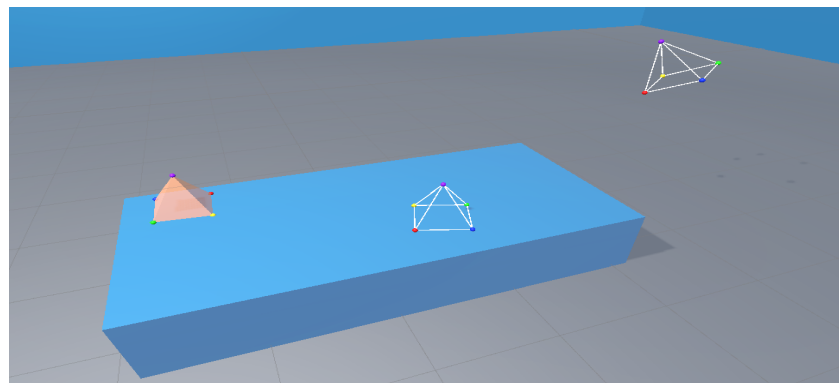
(**b**) Docking task.
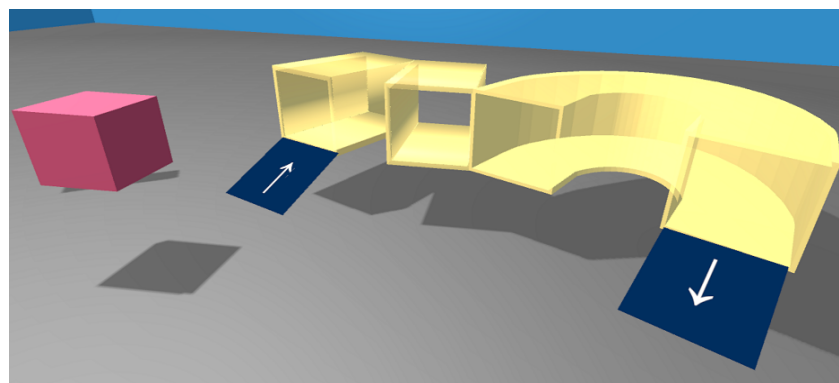
(**c**) Tunnel task.

Figure 5. Task models.

4.3. Procedure

The experiment began with the pre-experiment stage, where the participants filled out a demographic and Guilford–Zimmerman spatial ability questionnaire [43]. The pre-experiment stage was only conducted on the first day of the study when the participant arrived for the first time. After these questionnaires had been completed, the experiment entered the *training phase*, where we first introduced the technique through a demo video. After an explanation of the technique, we provided some simple tasks that participants needed to complete to acclimate to that condition. We provide instructions on how to accomplish the task using the interaction technique assigned under that condition. Then, they were allowed to practice repeatedly prior to the *testing phase*.

In the *testing phase*, participants began with a random task, as mentioned in Section 4.2. There was also a description of the task that was provided to the participants in the experiment environment. Once they understood the task, the participant clicked on a confirmation button. After the participants completed each trial, they had to deselect the

object and use a ray cursor to press the virtual 3D button to confirm that they were ready for the next trial. After clicking the confirmation button, the object for the next trial would appear immediately. The simulation gave audiovisual feedback to the confirmation button when clicking. The participants completed a total of 12 trials for all three tasks. Then, they completed a series of questionnaires in the *post-experiment* phase, including our self-created system performance questionnaire, the NASA-TLX Workload questionnaire [44], and the IPQ presence questionnaire [45].

In order to minimize or eliminate the effects of any carryover or learning between the three conditions, the participants returned approximately two days after each session to complete the other condition, in a manner similar to [33,37,46,47].

4.4. Measures

A number of measurements were collected in each trial in the study. The objective quantitative metrics consisted of movement time, number of attempts, number of collisions, position error, angular error, path length, and total rotation. A description of only the quantitative objective dependent variables that were statistically significant in our study is given below: Manipulation time: This is the time taken by the user to manipulate the object. The manipulation time starts when the user presses the trigger button to select, and ends when the user releases the trigger button. The mean manipulation time was used to measure how much time the user needed on average to translate or rotate the object of interest. The Number of Attempts: This is a measure that represents the number of times a user grabs and releases an object during each trial. The number of attempts is measured as the count of the number of times the user presses the trigger button to select and manipulate the object and then subsequently release it. Each time an object is selected to be manipulated and then subsequently released, the number of attempts per trail is incremented by one. The mean number of attempts is the average number of times users' select an object for manipulation and releases it across trials. The Number of Collisions: This is the number of times the manipulated object collided with other objects in the VR scene. The mean number of collisions shows on average how precisely and carefully users selected and manipulated the target object in the fine motor tasks across trials. The Angular Error: This is the sum of the absolute angular difference between the orientation of the target object and that of the reference object. Let $r_1 = (x_1, y_1, z_1)$ be the orientation of the reference object, and $t_2 = (x_2, y_2, z_2)$ be the orientation of the manipulated target object, with the orientation represented in Euler angles. The mean angular error is the average angular error of performance across trials in a task. The angular error AE is computed as

$$AE = |x_1 - x_2| + |y_1 - y_2| + |z_1 - z_2| \tag{1}$$

The objectives of this study were to compare these three interaction techniques using performance metrics of *efficiency*, the ability to quickly place the object in the destination; *accuracy*, the difference between the reference or ideal pose and the actual pose of the target, and the ability to place the object at the target without colliding with elements in the environment; and *economy of movement*, the ability of the user to manipulate the object directly to the target location without wasted or unnecessary movements. We quantified these three metrics using more specific variables that are listed above, where there is a many-to-one mapping between the objective quantitative variables and the performance metrics. Efficiency was quantified using the movement time and the number of attempts in each trial. Accuracy was quantified using the number of collisions, the distance error, and the angular error for each axis. Finally, economy of movement was quantified using the measures of path length and total angular rotation on each axis.

5. Results

5.1. Quantitative Objective Results

The objective variables were subjected to a one-way repeated measures ANOVA analysis, after verifying that all the assumptions of the parametric ANOVA analysis had been met (i.e., equality of variance, normality, and sphericity). The three within-subject conditions were Scaled HOMER, BMSR, and hybrid interaction techniques. The main goal of this was to determine how the user performance differed between each interaction technique. Pairwise post hoc tests between the levels of conditions were conducted using the Bonferroni method.

5.1.1. Pick-and-Place Task Performance

The ANOVA analysis found significant effects of the condition on the number of attempts ($F_{(2,54)}$ = 15.48, $p < 0.001$, part. η^2 = 0.41), number of collisions ($F_{(2,54)}$ = 4.29, p = 0.02, part. η^2 = 0.16), path length ($F_{(2,54)}$ = 5.33, p = 0.008, part. η^2 = 0.19), total rotation on the roll axis ($F_{(2,54)}$ = 3.26, p = 0.048, part. η^2 = 0.13), and angular error on the pitch axis ($F_{(2,54)}$ = 3.55, p = 0.024, part. η^2 = 0.15). Post hoc pairwise comparisons and illustrations of the magnitude of the significant differences using the Bonferroni method are shown in the graphs in Figure 6.
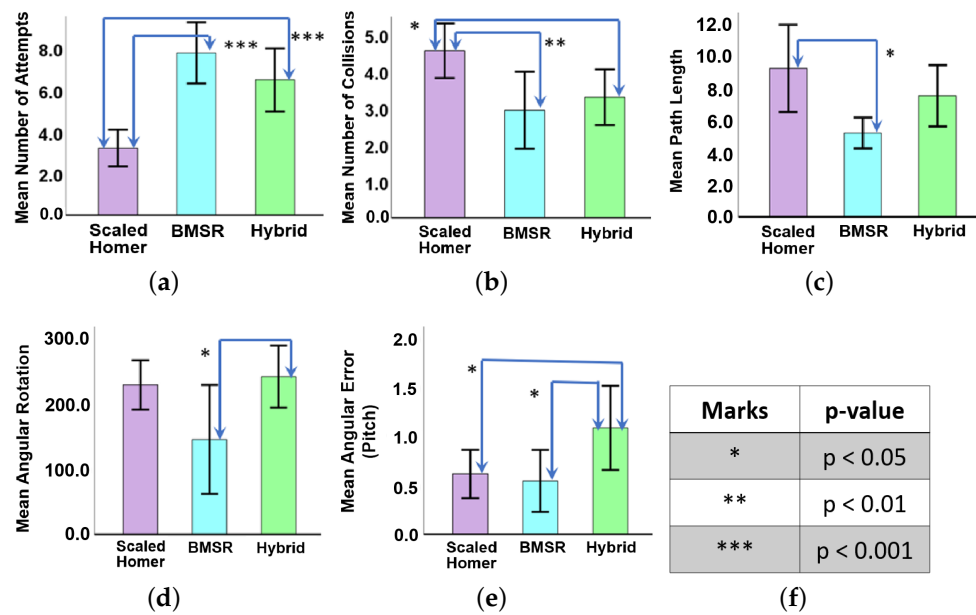

Figure 6. Boxplot Graphs showing the mean and confidence intervals of the significant effects of the condition in the pick-and-place task on (**a**) the number of attempts, (**b**) the number of collisions, (**c**) the path length, (**d**) the rotation on the roll axis, and (**e**) the angular error on the pitch axis. (**f**) The legend in the bottom right shows the strength of pairwise post hoc comparisons.

5.1.2. Docking Task Performance

The ANOVA analyses of the docking task showed significant effects of the condition on movement time ($F_{(2,34)}$ = 3.67, p = 0.033, part. η^2 = 0.14) and number of attempts ($F_{(2,34)}$ = 14.24, $p < 0.001$, part. η^2 = 0.39). Post hoc pairwise comparisons using the Bonferroni method are shown in the graphs in Figure 7a,b.
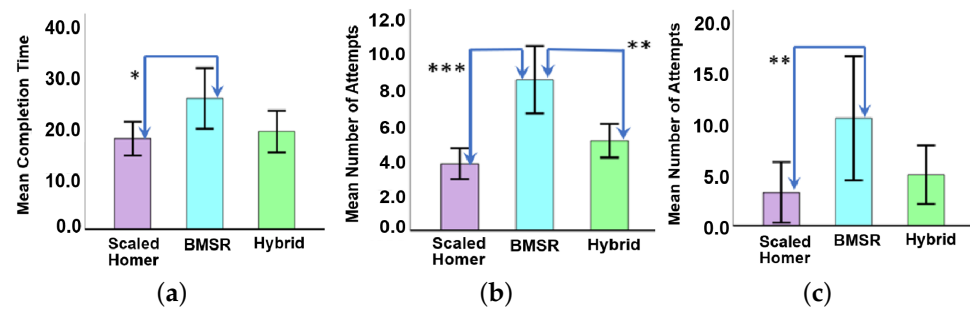
Figure 7. Boxplot Graphs showing mean and confidence intervals of the significant effects of the condition on (**a**) completion time in the docking task, (**b**) number of attempts in the docking task, and (**c**) number of attempts in the tunneling task. Strength of post hoc pairwise comparison is shown in the legend in Figure 6f.

5.1.3. Tunneling Task Performance

The ANOVA analyses of the data for the tunneling task showed a significant effect of the condition on the number of attempts ($F(2,34) = 3.32$, $p = 0.045$, part. $\eta^2 = 0.13$). Post hoc pairwise comparisons using the Bonferrroni method are shown in the graphs in Figure 7c.

5.1.4. Overall Performance Analysis

In order to examine the overall performance in all tasks, we pooled the data in all tasks and performed a repeated measures ANOVA analysis on the overall data (in a manner similar to previous analyzes), after verifying that all assumptions were met. ANOVA analysis revealed a significant main effect of condition on movement time ($F(2,34) = 4.300$, $p = 0.022$, part. $\eta^2 = 0.202$), on the number of attempts ($F(1.380,23.456) = 23.897$, $p < 0.001$, part. $\eta^2 = 0.584$), on the path length ($F(2,34) = 7.774$, $p = 0.002$, part. $\eta^2 = 0.314$), and in placement accuracy ($F(1.287,21.885) = 4.456$, $p = 0.038$, part. $\eta^2 = 0.208$). The graphs in Figure 8 show the results of post hoc pairwise comparisons using the Bonferroni method.
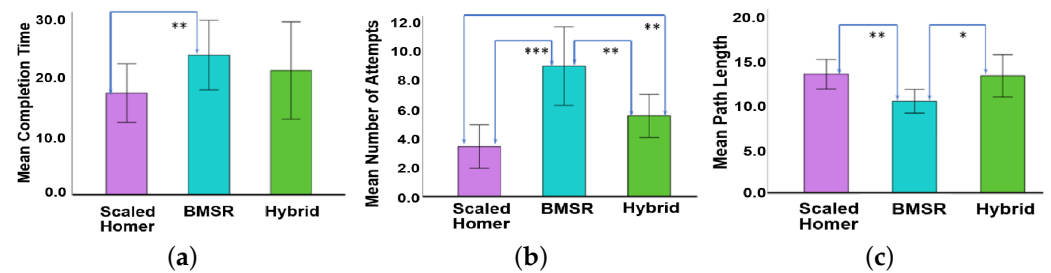

Figure 8. Boxplot Graphs of post-hoc analyzes showing the mean and confidence intervals of the significant effects of the condition on (**a**) completion time, (**b**) number of attempts, and (**c**) path length. The strength of the post-hoc pairwise comparison is shown in the legend in Figure 6f.

5.2. Quantitative Subjective Results

The subjective metrics were administered a non-parametric related-samples Friedman test, and we evaluated any significant effects via post hoc pairwise comparisons using Wilcoxon's signed ranks test.

5.2.1. System Performance Questionnaire

In evaluating the scores from our system performance questionnaire, we found the following results from the non-parametric analysis of the system performance questionnaire results.

In response to the question, "to what extent did you perceive you had sufficient motion control when moving an object from one location to another", we found that the condition significantly affected the perceived level of object motion control in moving the object via

translation $\chi^2 = 7.107$, $p = 0.029$. In the post hoc pairwise comparisons, Wilcoxon's signed ranks test revealed that the BMSR technique had a lower perceived translation control score than the hybrid technique (Z = −2.436, $p = 0.015$). See Figure 9a.

In response to the question, "to what extent did you perceive you had sufficient motion control in rotating an object", we found that condition also significantly affected the perceived level of object motion control in moving the object through rotation $\chi^2 = 19.433$, $p < 0.001$. Post hoc pairwise comparisons using Wilcoxon's signed ranks test revealed that the BMSR technique had a significantly lower perceived rotation control score than the hybrid technique (Z = −2.269, $p = 0.023$). Post hoc pairwise comparisons also revealed that Scaled HOMER had a lower perceived rotation control score than the BMSR technique (Z = −2.620, $p = 0.009$) and the hybrid technique (Z = −3.695, $p < 0.001$). See Figure 9b.

Finally, in response to the question, "to what extent did you perceive that you had sufficient motion control in moving an object from one location to another and rotating the object simultaneously", we found that condition also significantly affected the perceived level of object motion control in simultaneous translation and rotation $\chi^2 = 12.933$, $p = 0.002$. Post hoc pairwise comparisons using Wilcoxon's signed ranks test also revealed that the hybrid technique had a higher perceived simultaneous translation and rotation control score than Scaled HOMER (Z = −2.806, $p = 0.006$) and the BMSR technique (Z = −2.729, $p = 0.006$). See Figure 9c.
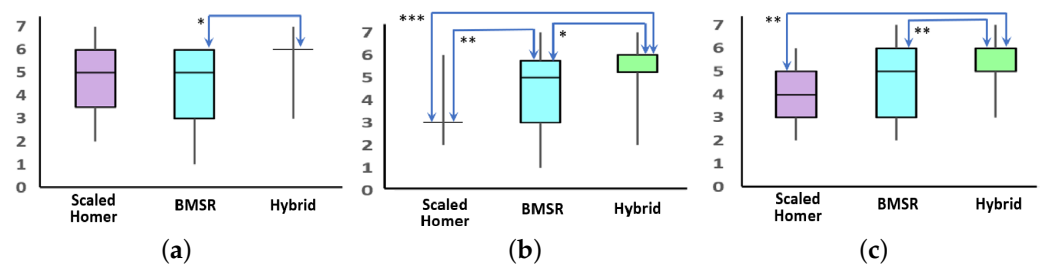

Figure 9. Boxplot graphs of post hoc analyses of the significant effects of the condition on perceived control of (**a**) translation alone, (**b**) rotation alone, and (**c**) translation and rotation simultaneously. Strength of post hoc pairwise comparison is shown in legend in Figure 6f.

5.2.2. NASA-TLX Workload Assessment

A non-parametric analysis revealed that the condition significantly affected the perceived mental demand $\chi^2 = 11.925$, $p = 0.003$, perceived physical demand $\chi^2 = 12.737$, $p = 0.002$, and perceived performance demand $\chi^2 = 8.291$, $p = 0.016$. Wilcoxon's signed ranks test revealed that the hybrid technique had a lower perceived mental demand than the Scaled HOMER (Z = −2.738, $p = 0.006$) and the BMSR techniques (Z = −2.949, $p = 0.003$). The signed ranks test revealed that the hybrid technique had a lower perceived physical demand than the Scaled HOMER (Z = −3.033, $p = 0.002$) and the BMSR techniques (Z = −2.992, $p = 0.003$). The signed ranks test revealed that the hybrid technique had a higher perceived performance demand than the Scaled HOMER (Z = −2.106, $p = 0.035$) and the BMSR technique (Z = −2.550, $p = 0.011$). These results are depicted in Figure 10.
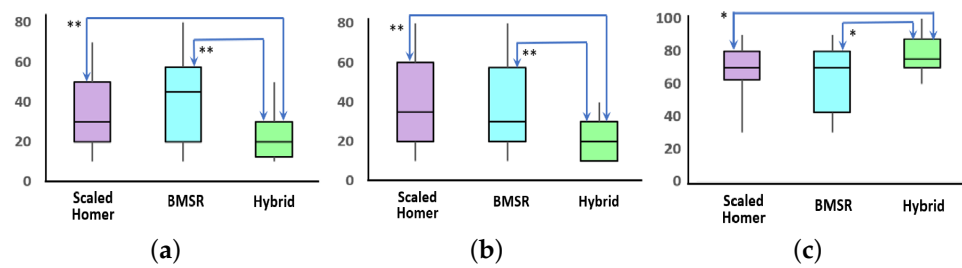
Figure 10. Boxplot graphs of post hoc analyses of the significant effects of the condition on (**a**) perceived mental demand, (**b**) perceived physical demand, and (**c**) perceived performance. Strength of post hoc pairwise comparison is shown in the legend in Figure 6f.

There was no significant difference in the affect of the condition on the presence scores.

5.3. Qualitative Results

As part of our system performance questionnaire, we asked each participant what they liked or disliked about each aspect of the simulation and which interaction metaphor they preferred of the ones that were available. When asked which metaphor they preferred between the Scaled HOMER and the near-field metaphor, the spread was relatively even. Out of 18 participants asked about this, 10 participants preferred the near-field metaphor and 8 participants preferred the Scaled HOMER metaphor. The responses to this question can be summarized by a participant who said "It depends. If you need to make long-range, fast movements, I prefer the Scaled HOMER; If you need to make large rotations or precise translations, I prefer the near-field metaphor". Two participants said that the "near-field replica blocked their view", and those who preferred the near-field metaphor said that it was easier to control, especially for more precise object placement.

When asked about what they liked or disliked about translational movements in the hybrid metaphor, seven participants stated that they liked that they could use Scaled HOMER for larger movements and the near-field metaphor to perform the fine-tuning. Similarly, when asked what they liked about rotations, six participants said that they could switch between scaled HOMER and near-field to harness the advantages of each. When asked specifically about their opinions about translation using Scaled HOMER, some of the participants stated that they liked the "intuitive control" and "extreme convenience for simple translations", and some stated that they did not like that it "required a lot of trial and error to get used to the relationship between the hand's velocity and the object's movement distance". When asked about their opinions about rotations, four participants said that they liked that they could rotate the object simply by rotating their wrists, but, similarly to translation, several participants said that they had a small range of angles to rotate and that it required a lot of trial and error to gain familiarity with the metaphor. When asked specifically about their opinions about translation in the near-field metaphor, some of the participants stated that they liked the fact that they could directly grab a replica of the object and interact with it, saying it "provided easier control" and "was realistic". However, six participants stated that they disliked the fact that they sometimes needed to select the same object multiple times. When asked about their opinions about rotations, five participants liked that they could perform precise rotations "due to the constrained transformation", and seven participants stated that they disliked that it was hard to decide which axis to select.

Finally, when asked which method they preferred between all three interaction metaphors, all but two of them preferred the hybrid metaphor, with most of them saying that it had the advantages of both the Scaled HOMER and the near-field metaphor techniques. One user preferred the near-field metaphor, and one preferred Scaled HOMER. The participant who preferred Scaled HOMER stated that it was "...more intuitive and faster". The participant who preferred the near-field metaphor responded that "...the replica

appears in front of the user eliciting more presence, can do direct manipulation using it. Also can not only do precise translation and rotation but also do intuitive manipulation, operation is more diverse". The questionnaire also asked about when they preferred to use each metaphor and why. When answering about Scaled HOMER, 10 participants stated that they preferred it when performing translational movements, specifically faster translations. Of those participants, six cited they liked the ability to perform "large-range translation". When participants were answering about when they preferred the near-field metaphor, 10 stated that they preferred it when performing rotational movements. In addition, six of these participants stated that they preferred it when performing precise movements.

6. Discussion

In order to answer our research question, "To what extent did participants' objective performance, subjective impressions and perceptions differ between the interaction techniques in the near field (BMSR with scaled replica), object space (Scaled HOMER) and a hybrid technique for object space or distant object manipulation in VR?", we first operationalized these research questions by formulating hypotheses to answer the underlying research question from an objective perspective. From Section 4, our first hypothesis (H1) was that the *BMSR technique would outperform Scaled HOMER in accuracy* and our second hypothesis (H2) was that the *BMSR technique would outperform Scaled HOMER in movement economy*. The first hypothesis was *supported* by our objective data, as the BMSR technique was shown to be superior in accuracy, based on mean angular error, mean angular rotation, and mean number of collisions, especially in the pick-and-place task. The examination of effect sizes also suggested that the effect of BMSR on accuracy over the other conditions was important and significant, as evidenced by the observed partial η square range of 0.14 to 0.21.

The second hypothesis was *partially supported* by our objective data. On the one hand, the economy of movement, based on mean path length, was superior for BMSR compared to Scaled HOMER in multiple tasks, such as pick-and-place and tunneling. On the other hand, the number of attempts for the stop and start of the hand movements in the manipulation process was less with Scaled HOMER compared to BMSR across all three tasks of pick-and-place, docking, and tunneling. Therefore, we found that H2 was partially supported overall. An examination of the effect sizes associated with economy of movement variables suggested that our results were also important and significant, as evidenced by the observed partial η square value range of 0.13 to 0.58.

We also hypothesized (H3) that *Scaled HOMER would result in quicker movement times than BMSR*. We found *support* for this hypothesis in our objective data in terms of mean number of attempts and speed, which were superior (i.e. lower) for Scaled HOMER compared to BMSR. An examination of the effect sizes associated with the efficiency variables suggested that our results were also important and significant, as evidenced by the observed partial η square value range of 0.14 to 0.20.

Our results suggest that the BMSR interaction technique offered better motion control than the Scaled HOMER or hybrid conditions, using the metric of fewer collisions, and greater movement economy, using the metric of lower path length, consistently across tasks. One possible reason for this finding could be the large distance over which users manipulated objects using Scaled HOMER and the instability of small hand motion that caused larger errors in placement, even with velocity-based scaling [1].

The BMSR interaction technique allows users to leverage near-field viewing, depth presentation, and perception, as well as visuo-proprioceptive information from hand/controller motion, which potentially enables precise control of objects, as supported by [17,31,32]. These cues provide maximum benefit when working with objects in a near-field space, which could provide an important advantage for near-field over object-space interaction techniques [48]. Additionally, the BMSR technique provides a scaled replica of the manipulated object, which potentially improves manipulation performance, as users can act on visuomotor information during fine motor actions on near-field replicas, as also shown by

research on the voodoo dolls interaction technique [27,49]. However, this may come at the cost of visibility of far-field objects, as the replica may partially occlude the participants' view.

Overall, we observed that the BMSR interaction technique had a lower motion instability than Scaled HOMER when manipulating distant objects. The BMSR technique allowed users to manipulate objects with degree-of-freedom (DoF) separation. This DoF separation can be beneficial for precise movements, as evidenced by the findings of studies by [4,5]. These studies showed that simultaneous movements of translation and rotation were desirable for long-range and faster movements, but separating the DoF was better for smaller and more precise movements. Participants made fewer attempts with the Scaled HOMER technique than in the BMSR technique. Specifically, in the docking task, the Scaled HOMER technique yielded lower task completion times than the BMSR technique, presumably reducing the time that was taken between each attempt.

Our findings are consistent with the results of the study by Katzakis et al., who found similar drawbacks for Scaled HOMER as in our findings [38]. These findings suggest that in terms of the speed–accuracy trade-off, participants tended to favor speed over accuracy with Scaled HOMER relative to the BMSR interaction technique. However, when using the BMSR technique, they tended more toward accuracy over speed [30]. This is a trade-off worth considering, especially since Scaled HOMER showed the same level of movement speed as other object-space manipulation techniques [1].

The fourth hypothesis (H4), which was that *the hybrid method would outperform Scaled HOMER and BMSR in accuracy and economy of movement*, was *not supported* by our objective data, as there were no significant performance differences in efficiency, accuracy, and economy of movement with the BMSR and Scaled HOMER techniques. Our hybrid technique allows the user to transition freely between Scaled HOMER and BMSR techniques. Although this transition was seamless, it still involves transitioning between methods, which could introduce additional dimensions of control and complexity for user interactions with this technique. However, interestingly, our hybrid technique had a lower perceived mental burden than the BMSR technique, as indicated by the NASA TLX workload results. Our participants even suggested that the hybrid interaction technique merged the best of both the constituent interaction techniques. All of these data could possibly imply a discrepancy between user impressions and objective performance, and this could suggest that interfaces that are perceived to be favorable may not always produce the best objective results [50]—what is favorable may not always be optimal.

The qualitative results provided additional support and clarified some of the objective quantitative findings. Participants' responses suggested that if they needed to make long-range (gross) motions, then they preferred Scaled HOMER for manipulation interactions. However, if they needed to make very precise translation or rotation manipulations, then they preferred the near-field interaction technique. The most interesting result was that, when asked which method they preferred between all three interaction techniques, the vast majority of participants preferred the hybrid interaction technique, as it took advantage of the *best of both words* approach, in that it integrated the advantages of Scaled HOMER for far field gross manipulation, and the near-field technique for personal space fine motor manipulations.

Limitations

Although there were not many objective differences between the hybrid technique and the other conditions, interestingly, our data revealed that the hybrid technique could yield similar objective results, while minimizing mental and physical demands compared to the Scaled HOMER or BMSR techniques individually. The hybrid technique aims to leverage BMSR and Scaled HOMER, thus we can expect that the user may use Scaled HOMER to rapidly and coarsely move the target object to a place near the destination and then use BMSR to manipulate the object into the final destination. If the tasks involve long-range translations, BMSR alone may require several rounds of object selection and manipulation,

and hence require more time than Scaled HOMER or the hybrid technique. In addition, although we found that the BMSR technique showed greater accuracy with regard to angular error compared to the other techniques, we expected to find more evidence of lower positional and orientation errors with the BMSR technique as compared to Scaled HOMER. We believe this remains to be explored further in future studies, where we can examine more specifically the fine motor actions of the two interaction techniques with manipulation tasks that require greater positional and orientation control, such as mechanical assembly or fine motor object extraction tasks. These tasks may also resemble concrete tasks in real applications of VR, compared to the abstract tasks typically used in basic interaction technique research. One finding that we noticed was that the analyses of the objective data yielded significant effects between conditions in the pick-and-place and docking tasks, but the tunneling task showed fewer significant differences between the conditions. This could be due to the added complexity of the tunneling task, as that task required a longer series of translating and rotating the object, and this could have caused a *ceiling effect* in performance between the three conditions, in that participants in all three conditions performed equally poorly in this task. Regarding the qualitative results, although we defined the preferences and user impressions questionnaire as neutral in language and tone, so as not to induce any bias in the manner in which the questions were asked, there could still have been some bias in the manner in which users responded to the questionnaire. However, we believe any such bias to be very small to non-existent, as the participants were not told what we expected in terms of the strengths and weaknesses of each condition, and our qualitative results also confirmed and validated the objective quantitative findings derived from the study.

7. Conclusions and Future Work

An advantage of virtual reality is that users do not need to be in direct proximity to the objects they are manipulating. This is helpful, as the user can take advantage of personal space depth presentation and motor control in manipulating far-field objects, without the need to approach the intended objects. In this paper, we set out to compare and contrast the effectiveness of a BMSR interaction technique with scaled replicas against Scaled HOMER, an established object-space or far-field manipulation technique, and our Hybrid interaction technique that combined both techniques via a seamless switching and transitioning method.

Our objective results suggested that Scaled HOMER yielded a faster performance than the BMSR technique, whereas the BMSR technique outperformed Scaled HOMER in the metrics of enhanced movement control (lowest collision) and economy of motion. This was reflected in the subjective information, as users preferred Scaled HOMER for fast movement but preferred the BMSR technique for fine control and adjustments. We also proposed a hybrid technique that allowed users to switch freely between the BMSR and Scaled HOMER techniques. Our data showed that, although there was no objective benefit to this hybrid technique as compared to the two constituent techniques, our subjective responses suggested that it was easier to use than the other two and reduced the overall perceived workload. The hybrid technique combined the advantages of both constituent techniques, but the added time and effort of switching between the two may counteract the benefits of its ease of use. Further research may reveal how a hybrid technique could yield objective benefits, to better reflect the users' subjective impressions.

Recommendations derived from this contribution include using a near-field interaction technique for indirect manipulation in tasks that require precise adjustments from a distance. However, object space manipulation techniques like Scaled HOMER may be better for larger translations and rotations. Near-field manipulation can be very useful in applications such as engineering, architecture, and mechanical assembly. Interface designers should consider the task they are trying to implement and then choose whether to use a near-field technique or a far-field object space manipulation technique, depending on how much precision is required to complete that task. Our findings have also shown that users prefer

a hybrid technique that can combine the precision of a near-field technique with the broad movements of a far-field manipulation technique. Our proposed hybrid technique was not shown to be objectively inferior to its component techniques.

A future direction of this research would be to explore the effects of DoF separation in near-field interaction techniques for improving precision and performance in manipulation tasks. We will also examine the effects of near-field, object-space, and hybrid interaction techniques on performance and perception in applied simulation scenarios such as mechanical assembly and fine object extraction.

Abbreviations

The following abbreviations are used in this manuscript:

DoF	Degree of Freedom
BMSR	Bimanual Near-field Metaphor with Scaled Replica
Scaled HOMER	Hand-centered Object Manipulation Extending Ray-casting

References

1. Wilkes, C.; Bowman, D.A. Advantages of velocity-based scaling for distant 3D manipulation. In Proceedings of the ACM Symposium on Virtual Reality Software and Technology, Bordeaux, France, 27–29 October 2008; pp. 23–29.
2. Frees, S.; Kessler, G.D. Precise and rapid interaction through scaled manipulation in immersive virtual environments. In Proceedings of the IEEE Conference on Virtual Reality, Bonn, Germany, 12–16 March 2005.
3. Frees, S.; Kessler, G.D.; Kay, E. PRISM interaction for enhancing control in immersive virtual environments. *ACM Trans.-Comput.-Hum. Interact. (TOCHI)* **2007**, *14*, *1*, pp. 2-es.
4. Veit, M.; Capobianco, A.; Bechmann, D. Influence of degrees of freedom's manipulation on performances during orientation tasks in virtual reality environments. In Proceedings of the 16th ACM Symposium on Virtual Reality Software and Technology, VRST '09, Kyoto, Japan, 18–20 November 2009; pp. 51–58. https://doi.org/10.1145/1643928.1643942.
5. Mendes, D.; Relvas, F.; Ferreira, A.; Jorge, J. The benefits of DOF separation in mid-air 3D object manipulation. In Proceedings of the ACM Conference on Virtual Reality Software and Technology, Munich, Germany, 2–4 November 2016; pp. 261–268.
6. Bowman, D.A.; Hodges, L.F. An evaluation of techniques for grabbing and manipulating remote objects in immersive virtual environments. In Proceedings of the Symposium on Interactive 3D Graphics, Providence, RI, USA, 27–30 April 1997; pp. 35–38.
7. Lee, C.Y.; Hsieh, W.A.; Brickler, D.; Babu, S.V.; Chuang, J.H. Design and empirical evaluation of a novel near-field interaction metaphor on distant object manipulation in VR. In Proceedings of the ACM Symposium on Spatial User Interaction, (SUI 2021), Virtual, 9–10 2021.
8. Napieralski, P.E.; Altenhoff, B.M.; Bertrand, J.W.; Long, L.O.; Babu, S.V.; Pagano, C.C.; Kern, J.; Davis, T.A. Near-field distance perception in real and virtual environments using both verbal and action responses. *ACM Trans. Appl. Percept. (TAP)* **2011**, *8*, 18.

9.   Bowman, D.; Kruijff, E.; LaViola, J.J., Jr.; Poupyrev, I.P. *3D User Interfaces: Theory and Practice*; Addison-Wesley: Boston, MA, USA, 2004.

10.   De Araújo, B.R.; Casiez, G.; Jorge, J.A.; Hachet, M. Mockup Builder: 3D modeling on and above the surface. *Comput. Graph.* **2013**, *37*, 165–178.

11.   Mapes, D.P.; Moshell, J.M. A two-handed interface for object manipulation in virtual reality. *Presence Teleoperators Virtual Environ.* **1995**, *4*, 403–416.

12.   Song, P.; Goh, W.B.; Hutama, W.; Fu, C.W.; Liu, X. A handle bar metaphor for virtual object manipulation with mid-air interaction. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Austin, TX, USA, 5–10 May 2012; pp. 1297–1306.

13.   Bossavit, B.; Marzo, A.; Ardaiz, O.; De Cerio, L.D.; Pina, A. Design choices and their implications for 3D mid-air manipulation techniques. *Presence Teleoperators Virtual Environ.* **2014**, *23*, 377–392.

14.   Cho, I.; Wartell, Z. Evaluation of a bimanual simultaneous 7DOF interaction technique in Virtual Environments. In Proceedings of the IEEE Symposium on 3D User Interfaces, Arles, France, 23–24 March 2015; pp. 133–136.

15.   Mendes, D.; Fonseca, F.; Araujo, B.; Ferreira, A.; Jorge, J. Mid-air interactions above stereoscopic interactive tables. In Proceedings of the IEEE Symposium on 3D User Interfaces (3DUI), Minneapolis, MN, USA, 29–30 March 2014; IEEE: Piscataway, NJ, USA, 2014; pp. 3–10.

16.   Gloumeau, P.C.; Stuerzlinger, W.; Han, J. PinNPivot: Object Manipulation using Pins in Immersive Virtual Environments. *IEEE Trans. Vis. Comput. Graph.* **2020**, *27*, 2488–2494. https://doi.org/10.1109/TVCG.2020.2987834.

17.   Mine, M.; Brooks, F.; Sequin, C. Moving objects in space: Exploiting proprioception in virtual environment interaction. In Proceedings of the SIGGRAPH 1997, Los Angeles, CA, USA, 3–8 August 1997; pp. 19–26.

18.   Mendes, D.; Caputo, F.M.; Giachetti, A.; Ferreira, A.; Jorge, J. A survey on 3D virtual object manipulation: From the desktop to immersive virtual environments. *Comput. Graph. Forum* **2019**, *38*, 21–45.

19.   Bowman, D.A.; McMahan, R.P.B.; Ragan, E.D. Questioning naturalism in 3D user interfaces. *Commun. ACM* **2012**, *55*, 78–88.

20.   Choi, M.G.; Lee, J.H.; Ha, W.; Lee, K.H. Optimal Close-up Views for Precise 3D Manipulation. *Comput. Animimation Virtual Worlds* **2019**, *30*, 78–88.

21.   Houde, S. Interactive design of an interface for easy 3D direct manipulation. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Monterey, CA, USA, 3–7 May 1992; pp. 135–142.

22.   Coĥe, A.; Décle, F.; Machet, M. tBox: A 3D transformation widget Designed for touch-screens. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Vancouver, BC, Canada, 7–12 May 2011; pp. 3005–3008.

23.   Poupyrev, I.; Billinghurst, M.; Weghorst, S.; Ichikawa, T. The Go-Go interaction technique: non-linear mapping for direct manipulation in VR. In Proceedings of the 9th Annual ACM Symposium on User Interface Software and Technology, Seattle, WA, USA, 6–8 November 1996; pp. 79–80.

24.   Mine, M. *ISAAC: A Virtual Environment Tool for the Interactive Construction of Virtual Worlds*; UNC Chapel Hill Computer Science Technical Report TR95-020; 1995. University of North Carolina, Chapel Hill, NC, USA.

25.   Liu, X.; Wang, L.; Luan, S.; Shi, X.; Liu, X. Distant object manipulation with adaptive gains in virtual reality. In Proceedings of the IEEE International Symposium on Mixed and Augmented Reality (ISMAR), Singapore, 17–21 October 2022.

26.   Stoakley, R.; Conway, M.J.; Pausch, R. Virtual reality on a WIM: interactive worlds in miniature. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Denver, CO, USA, 7–11 May 1995; pp. 265–272.

27.   Pierce, J.S.; Stearns, B.C.; Pausch, R. Voodoo Dolls: seamless interaction at multiple scales in virtual environments. In Proceedings of the Symposium on Interactive 3D Graphics, Atlanta, GA, USA, 26–29 April 1999; pp. 141–145.

28.   Guiard, Y. Asymmetric division of labor in human skilled bimanual action: The kinematic chain as a model. *J. Mot. Behav.* **1987**, *19*, 486–517.

29.   Lisle, L.; Lu, F.; Davari, S.; Tahmid, I.A.; Giovannelli, A.; Llo, C.; Pavanatto, L.; Zhang, L.; Schlueter, L.; Bowmanmendes, D.A. Clean the Ocean: An imersive VR experience proposing new modifications to Go-Go and WIM techniques. In Proceedings of the IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW), Christchurch, New Zealand, 12–16 March 2022.

30.   MacKenzie, I.S.; Isokoski, P. Fitts' throughput and the speed-accuracy tradeoff. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Florence, Italy, 5–10 April 2008; pp. 1633–1636.

31.   Arsenault, R.; Ware, C. The importance of stereo and eye-coupled perspective for eye-hand coordination in fish tank VR. *Presence Teleoperators Virtual Environ.* **2004**, *13*, 549–559.

32.   Brickler, D.; Volonte, M.; Bertrand, J.W.; Duchowski, A.T.; Babu, S.V. Effects of stereoscopic viewing and haptic feedback, sensory-motor congruence and calibration on near-field fine motor perception-action coordination in virtual reality. In Proceedings of the 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), Osaka, Japan, 23–27 March 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 28–37.

33.   Ebrahimi, E. Investigating Embodied Interaction in Near-Field Perception-Action Re-Calibration on Performance in Immersive Virtual Environments. Ph.D. Thesis, Clemson University, Clemson, SC, USA, 2017.

34.   Bhargava, A.; Bertrand, J.W.; Gramopadhye, A.K.; Madathil, K.C.; Babu, S.V. Evaluating multiple levels of an interaction fidelity continuum on performance and learning in near-field training simulations. *IEEE Trans. Vis. Comput. Graph.* **2018**, *24*, 1418–1427.

35. Arora, R.; Habib Kazi, R.; Grossman, T.; Fitzmaurice, G.; Singh, K. Symbiosissketch: Combining 2D & 3D sketching for designing detailed 3D objects in situ. In Proceedings of the CHI Conference on Human Factors in Computing Systems, Montreal, QC, Canada, 21–26 April 2018; pp. 1–15.

36. Montano-Murillo, R.A.; Nguyen, C.; Kazi, R.H.; Subramanian, S.; DiVerdi, S.; Martinez-Plasencia, D. Slicing-Volume: Hybrid 3D/2D multi-target selection technique for dense virtual environments. In Proceedings of the 2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), Atlanta, GA, USA, 22–26 March 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 53–62.

37. Ebrahimi, E.; Altenhoff, B.M.; Pagano, C.C.; Babu, S.V. Carryover effects of calibration to visual and proprioceptive information on near field distance judgments in 3D user interaction. In Proceedings of the 2015 IEEE Symposium on 3D User Interfaces (3DUI), Arles, France, 23–24 March 2015; IEEE: Piscataway, NJ, USA, 2015; pp. 97–104.

38. Katzakis, N.; Seki, K.; Kiyokawa, K.; Takemura, H. Mesh-Grab and Arcball-3D: Ray-based 6-DOF object manipulation. In Proceedings of the 11th Asia Pacific Conference on Computer Human Interaction, Bangalore, India, 24–27 September 2013; pp. 129–136.

39. Kulik, A.; Kunert, A.; Froehlich, B. On motor performance in virtual 3D object manipulation. *IEEE Trans. Vis. Comput. Graph.* **2020**, *26*, 2041–2050.

40. Louis, T.; Berard, F. Superiority of a handheld perspective-coupled display in isomorphic docking performances. In Proceedings of the ACM International Conference on Interactive Surfaces and Spaces, Brighton, UK, 17–20 October 2017; pp. 72–81.

41. Grandi, J.G.; Berndt, I.; Debarba, H.G.; Nedel, L.; Maciel, A. Collaborative 3D manipulation using mobile phones. In Proceedings of the 2016 IEEE Symposium on 3D User Interfaces (3DUI), Greenville, SC, USA, 19–20 March 2016; pp. 279–280. https://doi.org/10.1109/3DUI.2016.7460079.

42. Thomas, B.H.; Lindeman, R.; Marchal, M. *Symposium Chair Message*; IEEE Symposium on 3D User Interfaces, Greenville, SC, USA, 19–20 March 2016.

43. Guilford, J.P.; Zimmerman, W.S. The Guilford-Zimmerman aptitude survey. *J. Appl. Psychol.* **1948**, *32*, 24.

44. Hart, S.G.; Staveland, L.E. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in Psychology*; Elsevier: Amsterdam, The Netherlands, 1988; Volume 52, pp. 139–183.

45. Schubert, T.; Friedmann, F.; Regenbrecht, H. Decomposing the sense of presence: Factor analytic insights. In Proceedings of the 2nd International Workshop on Presence, Colchester, UK, 6–7 April 1999; Volume 1999.

46. Lin, Y.X.; Venkatakrishnan, R.; Venkatakrishnan, R.; Ebrahimi, E.; Lin, W.C.; Babu, S.V. How the presence and size of static peripheral blur affects cybersickness in virtual reality. *ACM Trans. Appl. Percept. (TAP)* **2020**, *17*, 1–18.

47. Day, B.; Ebrahimi, E.; Hartman, L.S.; Pagano, C.C.; Robb, A.C.; Babu, S.V. Examining the effects of altered avatars on perception-action in virtual reality. *J. Exp. Psychol. Appl.* **2019**, *25*, 1.

48. McIntire, J.P.; Havig, P.R.; Geiselman, E.E. Stereoscopic 3D displays and human performance: A comprehensive review. *Displays* **2014**, *35*, 18–26.

49. Pierce, J.S.; Pausch, R. Comparing Voodoo Dolls and HOMER: Exploring the importance of feedback in virtual environments. In Proceedings of the CHI Conference on Human Factors in Computing Systems, Minneapolis, MN, USA, 20–25 April 2002; pp. 105–112.

50. Lubos, P.; Bruder, G.; Steinicke, F. Analysis of direct selection in head-mounted display environments. In Proceedings of the 2014 IEEE Symposium on 3D User Interfaces (3DUI), Minneapolis, MN, USA, 29–30 March 2014; pp. 11–18.