Characterizing Asymmetric Collaborative Interactions in Virtual and Augmented Realities

Jerônimo Gustavo Grandi*[1][2] Henrique Galvan Debarba[3] Anderson Maciel[1]

[1] Institute of Informatics, Federal University of Rio Grande do Sul, Porto Alegre, Brazil
[2] Mechanical Engineering and Materials Science, Duke University, Durham, United States
[3] IT University of Copenhagen, Copenhagen, Denmark

ABSTRACT

We present an assessment of asymmetric interactions in Collaborative Virtual Environments (CVEs). In our asymmetric setup, two co-located users interact with virtual 3D objects, one in immersive Virtual Reality (VR) and the other in mobile Augmented Reality (AR). We conducted a study with 36 participants to evaluate performance and collaboration aspects of pair work, and compare it with two symmetric scenarios, either with both users in immersive VR or mobile AR. To perform this experiment, we adopt a collaborative AR manipulation technique from literature and develop and evaluate a VR manipulation technique of our own. Our results indicate that pairs in asymmetric VR-AR achieved significantly better performance than the AR symmetric condition, and similar performance to VR symmetric. Regardless of the condition, pairs had similar work participation indicating a high cooperation level even when there is a visualization and interaction asymmetry between the participants.

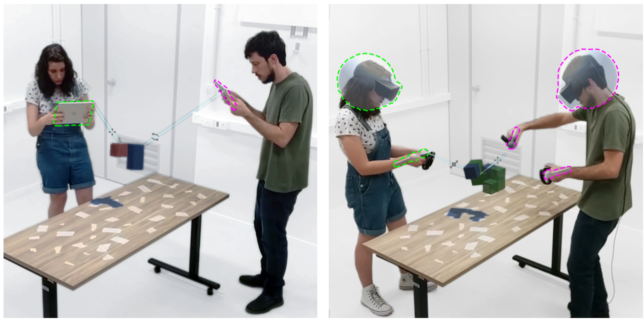
Index Terms: Human-centered computing—Human computer interaction (HCI)—Interaction techniques; Human-centered computing—Human computer interaction (HCI)—Interaction paradigms—Mixed/augmented reality Human-centered computing—Collaborative and social computing

1 INTRODUCTION

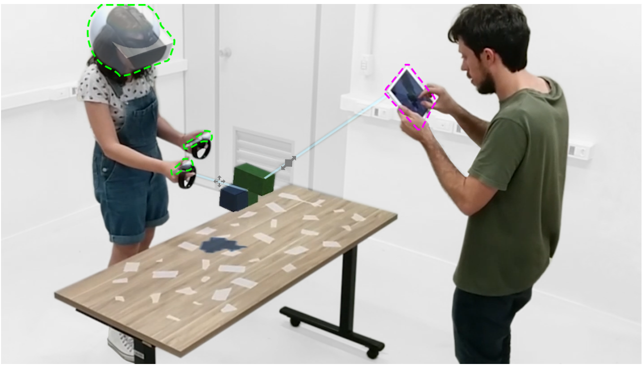
Collaborative Virtual Environments (CVEs) are shared spaces designed to support interactions between users and objects in Virtual Reality (VR) [9]. CVEs add the possibility to integrate different visualization and interaction interfaces into a mixed reality environment, where users can co-exist in an asymmetric experience. In the context of CVEs, asymmetry represents the capacity of individuals in a group to have different means to visualize and interact with virtual content. That is, collaborating users interact through completely different sensorimotor configurations and build distinct perspectives of the shared experience.

The simultaneous manipulation (i.e. translation, rotation and scale) of the same object by multiple users, also called co-manipulation, can enhance the team capability to solve complex manipulation tasks, such as the accurate positioning of an object, when compared to the non-collaborative setting [12]. Teamwork involves considerable negotiations and, as team members vary, team strategies and task-fulfillment processes also change. Modeling such interactions raises novel collaborative concepts compared to those typically grounded in a single-user scenario [29]. What is not clear is whether these patterns of collaboration emerge in collaborative asymmetric scenario, which might enforce different individual strategies to each user.

In this paper, we present an assessment of asymmetric interactions in Collaborative Virtual Environments. In our asymmetric setup, two co-located users interact with virtual 3D objects, one in immersive Virtual Reality (VR) and the other in mobile Augmented Reality (AR), as shown in Fig. 1. While individual interactions in asymmetric Mixed Reality (MR) environments have been explored in the past (see Sect. 2.1), this work investigates the effectiveness of simultaneous and cooperative object manipulation. We compared two *Symmetric* setups, where the interaction and visualization interfaces are the same for both participants (i.e. both interact in VR or both interact in AR), with the *Asymmetric* setup, where interaction and visualization are distinct for each participant (i.e. one user interacts in VR and the other interacts in AR). In the assessment we measured the impact that asymmetric interactions have on performance, social interaction and awareness.



(a) Symmetric AR (b) Symmetric VR



(c) Asymmetric VR and AR

Figure 1: Two users simultaneously interacting with a 3D object. Participants share the same physical space regardless of the interface. The techniques allow for manipulations beyond the arm's reach. Rays drawn from the hand to the object represent the selection and icons inform the actions. The dashed outlines represent the user's avatars in the VE. All virtual elements are synchronized among the participants.

To allow for fluid simultaneous manipulation of 3D objects by

---

*e-mail: jggrandi@gmail.com

several users we rely on the approach that we described in [12] to handle simultaneous inputs from multiple participants. For the AR setup, we choose the collaborative AR interface proposed in [13]. For the VR setup, we observed a lack of VR solutions that conforms to the *Sum-of-Contributions* [12] approach and are suitable to current consumer VR equipment. As a result, we designed a novel collaborative 3D user interface for virtual object manipulations relying on a pair of 6 degrees of freedom (DOF) controllers. In addition, we also perform an assessment of the proposed VR interface by comparing it to traditional direct object manipulation in VR, which relies on the virtual hand metaphor (i.e. grab interpenetrating objects and manipulate with hand centered translation and rotation).

In summary, the contribution of this paper is twofold: (i) the design and assessment of a collaborative 3D manipulation interface for immersive VR; and (ii) the assessment of collaborative asymmetric interactions, with insights on the impact of asymmetry in social interaction and awareness.

The paper is organized as follows. Sect. 2 recalls the ralated work on collaborative and asymmetric interaction. Then, the collaborative VR co-manipulation technique is described and tested in Sect. 3.2 and Sect. 5 respectivelly. In Sect. 6, we investigate the effect of an asymmetric setup on co-manipulations. Results are discussed in Sect. 7. Finally, the general conclusion in Sect. 8 ends the paper.

2 RELATED WORK

2.1 Asymmetric Interaction in Collaborative Mixed Reality Environments

Generally, asymmetry in VEs can vary in scale, Point-of-View (PoV), roles, realities and devices, and depending on the design they are linked together. One way to explore asymmetry is to provide participants with different scales in the VE, and consequently, allowing for different Points-of-View. Asymmetric scales are assigned to different users and explored in navigation [24, 38], manipulation of 3D objects [8], data exploration [36], architectural design [11, 15] and in sports [39].

While asymmetry in scale is well explored, CVEs can be also designed to handle asymmetric interactions while users keep the same scale. Those systems usually support interaction with different input devices and visualization hardware while preserving its usability across the asymmetric modalities. ShareVR [14] combines *HMD* users with *Non-HMD* users into a spatial augmented reality and virtual reality environment. The prototype enables co-located users to interact with each other in the same shared environment. The authors report an increase in enjoyment, presence and social interaction when compared with the *Non-HMD* user interacting in a conventional TV+Gamepad setup. In the same direction, CoVAR [26] implements collaborative interactions between remote AR and VR systems. The AR user scans the local environment and shares it with the VR user to increase awareness of the space during the interactions. Roo et al. [30] proposed the *One Reality*, a hybrid mixed reality conceptual framework that allows multiple users to coexist simultaneously and to transition between multiple mixed reality modalities while interacting with virtual objects. Asymmetric PoVs are commonly explored for collaborative interactions in 360 degree videos [23, 17]

The popularity of HMDs brought the asymmetric metaphor to commercial games, such as, in the *Playroom VR*[1] *and Black Hat Cooperative*[2], where a user wears the HMD and has a different perspective from players who are using the conventional game controllers and screen.

Similarly to prior works, we allow users to cooperate using asymmetric interactions. However, in our solution users can share the 3D transformation tasks in a fluid simultaneous manipulation while keeping the same scale and sharing the same physical space.

2.2 Co-manipulation of 3D Objects

The work of Margery et al. [19] classifies the collaborative interactions based on the features provided to the user for cooperation. According to their classification, simultaneous manipulations have the highest cooperation level. Different from the real world, where a physical link is created between the users that are carrying an object, in the virtual world this link does not exists. Thus, techniques that use a tangible user interface (TUI) [1, 32] seem to be the natural choice to provide a physical form to control virtual objects. Even though TUIs can increase task performance and collaboration due to the direct mapping of movements and the tangibility of the props, TUIs are fundamentally designed for specific applications [32] and they are rather limited by the shape of the prop that can only represent a set of virtual objects even in the configurable approach [1].

On the other hand, without a TUI, the management of concurrent access to the same object by multiple users is necessary to achieve efficient co-manipulations. Ruddle et al. [31] classify the techniques into the symmetric integration of movements, where users perform coordinated actions (e.g. both users move the object up), and asymmetric integration of movements, where users can perform different actions on the object (e.g. one user rotates another user moves). The authors compare the symmetric and asymmetric approaches in a study where users need to carry a large object through a corridor. They report that the symmetric approach is more relevant when the two users have to perform a very similar action, while the asymmetric approach is preferred when users have to perform different tasks. The Bent-Pick-Ray [28] is an asymmetric technique for co-manipulation in VR based on the ray-casting metaphor. When two users are manipulating the same object, the selection rays bent based on the pointing direction and the point of selection. This visual feedback helps users understand each other's actions since manipulations are simultaneous. Pinho et al. [25] explore the Degrees-of-Freedom (DOF) separation to demonstrate that the use of cooperative interaction techniques can be more efficient than two users working in parallel using single-user interaction. In this case, the number of DOFs that each user can access and control is limited: one user controls the rotation of the object, while the other one is limited to translation. The SkeweR [10] technique enables multiple users to simultaneously grab any part of a virtual object through points called "crushing points". To determine the translation and the rotation of a grabbed object, SkeweR considers positions of these points and average the 3D object final position.

To make collaboration in VEs more accessible to multiple users, we designed a handheld-based interface for collaborative object manipulation for shared displays [12]. In that previous work, we use regular smartphones as a 3D user interface to collaboratively control virtual objects on a shared screen. We then compared the performance of different group sizes during synchronous 3D manipulation tasks. The technique was extended for AR handheld devices in [13]. That interface design was based on the techniques that combine both the device movements and touch gestures for 3D manipulations [22, 20]. In this design, several users can interact through their point-of-view and simultaneously manipulate 3D virtual augmentations in the same shared physical space. In the present paper, we expand research on co-manipulation of 3D Objects by evaluating the effect of asymmetric CVEs.

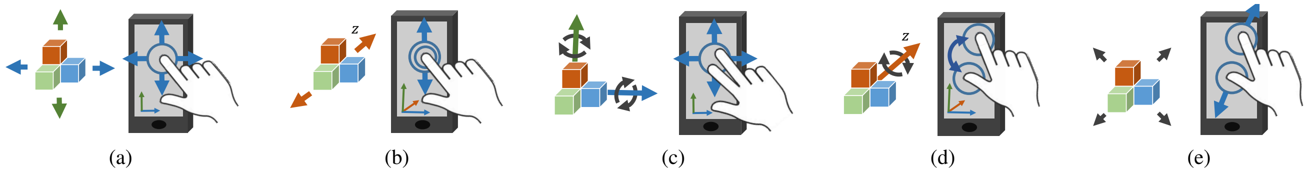
3 3D CO-MANIPULATION INTERFACES

Here we describe the 3D user interfaces we used for collaborative manipulations to evaluate collaborative asymmetric interactions in AR and VR environments.

---

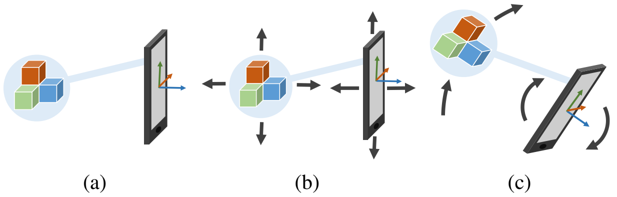
[1] https://www.playstation.com/en-us/games/the-playroom-vr-ps4/

[2] https://www.teamfuturegames.com/



(a) (b) (c) (d) (e)

Figure 2: Manipulations performed over the selected object: (a) A touch and slide translates in xy axis, (b) one tap followed by a touch and slide translates in z axis, (c) touch and slide with two fingers rotates in xy axis, (d) touch and rotate two fingers rotates in z axis and (e) pinch and spread of two fingers to uniformly modify the object's scale.



(a) (b) (c)

Figure 3: Manipulations performed on the attached object: (a) object-device attachment, the object and the device keep an invariant rigid transformation. In this way, the object translates (b) and rotates (c) with the device movements.

3.1 Augmented Reality Interaction

We choose the Augmented Reality interface we proposed in a previous work [13] to handle the interactions using handheld devices. The interface demonstrated to the effective for co-manipulation in collaborative environments. It integrates two input possibilities for 3D transformations: one using touch gestures (Fig. 2) and the other using the device movements, where the object is attached to the physical pose of the device (Fig. 3).

Moreover, it uses the *Sum-of-Contributions* approach [12] in the design of the collaborative manipulations. The *Sum-of-Contributions* seamlessly aggregate inputs from multiple sources directly and unrestrictedly.

3.2 Immersive Virtual Reality Interaction

The *Sum-of-Contributions* approach is a simple yet robust way to integrate simultaneous manipulation of virtual objects independently of the device in use. Thus, we implemented it in our VR technique to handle co-manipulation.

3.2.1 Approach

Interaction in immersive VR typically uses the *Virtual Hand* metaphor [16] to directly map the user's hand movements to their virtual hand. While the *Virtual Hand* is a natural way to interact with 3D objects, the interaction space is constrained by the arms' reach. Thus, selecting and manipulating objects out of reach usually requires user re-positioning, which can be achieved by walking or through a navigation technique. Moreover, the egocentric nature of grasping methods inhibits the simultaneous manipulation of objects by multiple users. To cope with these constraints, we designed a co-manipulation interface for translation, rotation and uniform scaling of objects beyond arms reach. The interactions may occur with one hand or combining both hands with synchronous movements. Our design consists of a variation of the *HOMER (Hand-centered Object Manipulation Extending Ray-Casting)* technique [5] for single-handed manipulation and of the *Spindle* technique [18] for two-handed manipulation. Moreover, our technique handles multiple concurrent and synchronous inputs for the same DoF by an unlimited number of users (Level 3.2 on the Margery et al. [19] classification). The technique is implemented to work on commercial HMDs controllers, such as the Oculus Rift and the HTC Vive.

3.2.2 3D Manipulations

A selected object is handled regardless of the distance that it is from the user's virtual hand. The *trigger* button is used to select the object. In single hand manipulations, while the button is pressed, translation and rotation are applied directly with a 1:1 mapping. As a result, manipulation of objects within reach works similarly to what occurs in the *Virtual Hand* technique (Fig. 4a). Besides, with our technique, users can repeat the movement to move the object away or closer or to reach a total rotation beyond the wrist limits (i.e. press the trigger, move, release the trigger, position the hand, repeat). Finally, the up and down directions in the joystick uniformly modify the object's scale. The scale increments can vary depending on the position of the joystick.

Bimanual interaction is activated whenever the *trigger* button is pressed and held in both hands. In this mode, the hands perform symmetric-synchronous movements to execute the transformations [37]. The move of the two hands in the same direction translates the object (Fig. 5a). Moving the hands in opposite directions rotates the object in the y (one hand forward and other backward) and z (one hand upward and other downward) axis (Fig. 5b). Rotation around the x axis is achieved by rotating both hands around that axis, which differentiates our technique from *Spindle* technique [18]. Finally, moving the hands away or closer to each other changes the object scale (Fig. 5c).

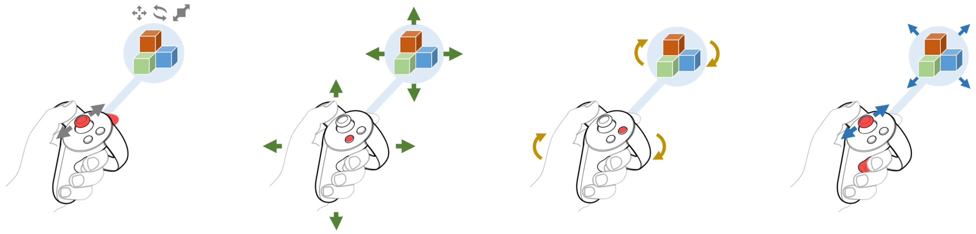
All three transformations can be executed simultaneously in both single and bimanual interaction. Additionally, it is possible to pick a transformation and block the others. Fig. 4 shows the actions to lock specific transformations.The transformation locks are optional, and can be combined by pressing more than one button at the same time. Moreover, the *trigger* button has priority over the lock buttons, when pressed while other lock button is pressed, it unlocks all transformations. The transformations lock resource adds more strategy possibilities during simultaneous manipulations. For instance, the groups can define roles during the manipulation task in order to separate the DOFs. We added virtual avatars to make the user aware of other collaborating users. Every time the user manipulates the object, we draw virtual rays connecting the interacting hand with the object. Iconic representations inform which DOFs the user is controlling (see Fig. 1). These visual feedbacks are synchronized so every participant knows each others selections and actions.

4 EVALUATION DESIGN

The VR interface and asymmetric evaluations, described in the following sections, share the same task parameters and apparatus setup.

4.1 Task

We designed a 3D docking task for our experimental setup. The docking consists of aligning the position (3 DOFs), rotation (3



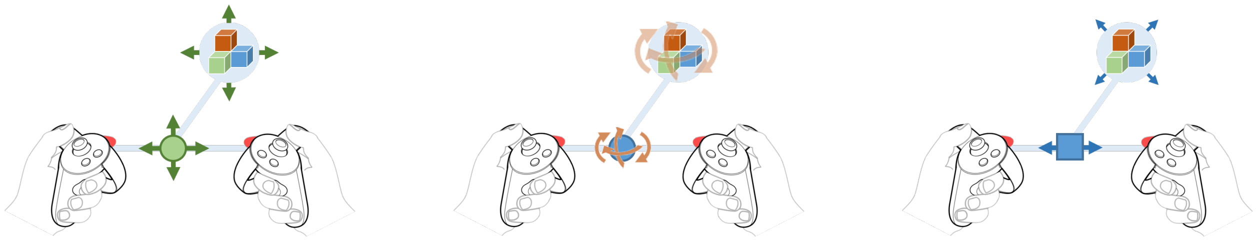
(a) Integrated Translation, Rotation and Scale

(b) Translation Lock

(c) Rotation Lock

(d) Scale Lock

Figure 4: Interaction gestures for distant manipulations with one hand. (a) While pressing and holding the trigger button with the index finger it is possible to translate, rotate. While holding the object, moving the joystick in up and down directions uniformly modify the object's scale. The three transformations can be combined and performed simultaneously. Transformations locking allow to focus on only one transformation at a time. It is achieved by pressing specific buttons on the controller interface. It is possible to individually alter the translation (a), the rotation (b) and the scale (c). It is possible to lock more than one transformation by pressing more than one lock button at the same time.



(a) Bimanual Translation

(b) Bimanual Rotation

(c) Bimanual Scale

Figure 5: Bimanual transformations inspired on the work of Mapes and Moshel [18].

DOFs) and scale (1 DOF, uniform scale) of the controlled virtual object with an identical *target* object. The target is set to 50% opacity to facilitate visualization when both the controlled and the *target* objects interpenetrate. Cube blocks similar to the Shepard and Metzler [35] construction compose the virtual objects. The blocks have 6.5cm long edges and have different colors to avoid ambiguity. The object is composed by the maximum of two blocks aligned, totaling 13cm in size on each axis. The *target* has fixed size, while the controlled objects have four possible initial scales: 60%, 80%, 120%, 140% of the *target* size. The maximum threshold tolerance for a successful docking was set to: **1cm** in position, **3°** in rotation and **1% difference** in scale.

We recreated the physical room in the virtual world. Both physical and virtual spaces are aligned. This provides passive haptics that allow the participant to touch and feel the walls and the table. Moreover, in the collaborative evaluation, co-located users coexist in the same shared VE and, therefore, can communicate with accurate positional sound. A manual calibration procedure is only necessary for the VR setup. It consists in positioning the participant in a specific position in the physical space and recenter the VE. In AR, the calibration step is not necessary since fiducial markers provide a reference for alignment.

4.2 Apparatus

The apparatus for visualization and interaction is composed of a pair of VR headsets and a pair of AR handheld devices. The VR technique evaluation only required one VR headset. The VR headset used is the *Oculus Rift* (1080×1200 pixels per eye, 90Hz refresh rate, $\approx 100^\circ$ field of view) with *Touch Controllers*. Each headset had its own tracking sensors and was connected to its own PC. Both machines had similar configuration. The tracking sensors were positioned to avoid occlusions by the other subject. The AR device is an Apple iPad Air 2 with 9.7 inches screen (264ppi density, $\approx 69.9^\circ$ diagonal camera field of view) and weights 437g. The AR tracking is performed by the Vuforia SDK. We track the mobile device physical pose relative to a fiducial marker. We used the table's surface as the reference marker and the extended tracking feature to expand the interaction range. Besides the interaction hardware, we used a dedicated server to manage the experimental parameters and to record all user interactions during the evaluation. All devices were connected to a dedicated network with wired and WiFi connection. Both server and client's application were developed using the Unity3D game engine and Unity's UNET API handles the network communication.

5 VR TECHNIQUE EVALUATION

We first conducted an experiment to assess the performance of the *Distant* technique design compared to the more widely used *Virtual Hand* metaphor in a single user setting. The *Virtual Hand* is a 3D interface metaphor that mimics the grasping and posing of a real object. It is natural and easy to understand and, as a result, widely adopted in VR applications.

It seems reasonable to assume that the *Virtual Hand* is the "gold standard" and, consequently, will perform better. Thus, in this experiment, we intend to assess the performance gap —particularly the completion time— between our *Distant* technique and the *Virtual Hand* techniques.

5.1 Procedure

The experiment was conducted using a within-subjects design with the *Technique (Virtual Hand, Distant)* as the sole independent variable of this evaluation. The presentation order of the conditions was counterbalanced. We collected the *time* to complete each docking,

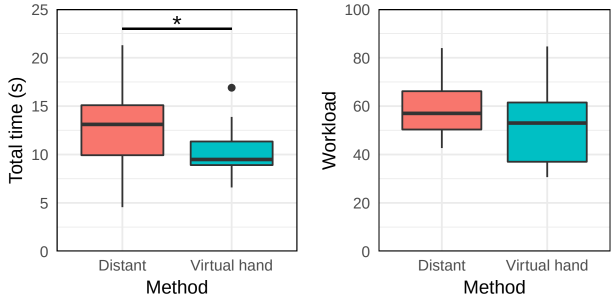


Figure 6: Box and whiskers plot of the trial completion time and workload for each level of the Technique variable. Users perform faster with the *Virtual Hand* technique ($M = 10.32$, $SD = 2.95$) than with the *Distant* technique ($M = 12.69$, $SD = 4.32$). No statistically significant difference was found for workload.

participants *workload* (NASA TLX questionnaire [3]), and participants subjective evaluation of task difficulty with a Simple Easy Question (SEQ) [33] as dependent variables.

After a training session, that consisted in a guided demonstration and three practice docking trials, the participants had to complete eight docking trials in sequence. The trials' presentation order was randomized. We instructed participants to complete each trial as fast as possible but avoiding to rush, that could lead to inaccuracy. After each condition, while resting, the subjects answered the Single Easy Question (SEQ) [33] and the NASA TLX questionnaire. Then, this procedure was repeated for the second experimental condition. At the end of the experiment, participants answered System Usability Score (SUS) [7] questionnaire to assess the overall usability of the experimental setup.

We analyzed whether the independent variable had a significant effect on the *trial time* with a repeated measures ANOVA test. We verified if the ANOVA assumption of normality of the residuals was violated with the Shapiro-Wilk test. The Wilcoxon signed-rank test was used for the NASA TLX and the SEQ results. The alpha significance level was set to 0.05 in all cases. A total of twelve participants with none or very little experience with HMD VR voluntarily took part on this experiment. In summary, the experiment design had: **12** subjects × **2** conditions × **8** trials = **192** unique trials.

5.2 Results

5.2.1 Virtual Hand vs. Distant Manipulations

We found a significant effect for the interaction *Technique* variable on completion time ($F_{1,11} = 5.68$, $p = .036$), with *Virtual Hand* performing faster than *Distant* manipulation. Results of the NASA TLX questionnaire did not yield a statistically significant difference between the two techniques ($p = .108$). These results can be visualized in Fig. 6.

The SEQ consisted of the statement "Overall, this task was ...", with the answer given on a seven point scale with 1 for "very hard" and 7 for "very easy". Participants found the task easier with the *Virtual Hand* technique ($Mdn = 6$) than with the *Distant* technique ($Mdn = 5$, $p = .019$).

Finally, the System Usability Score (SUS) ranged from 62.5 to 95 ($M = 80$, $SD = 9.36$). According to surveys that compare SUS scores for different systems, the system score is ranked as ”Good” [3].

6 ASSESSMENT OF COLLABORATIVE ASYMMETRIC INTERACTIONS

Here we investigate the effectiveness of simultaneous manipulation when both users are manipulating the same 3D object. We conduct an experiment to explore collaborative asymmetric interactions in a shared VR and AR environment when two users are co-manipulating virtual objects. In our asymmetric setup, two co-located users interact with virtual 3D objects in immersive Virtual Reality (VR) and mobile Augmented Reality (AR), as shown in Fig. 1. We want to assess the impact that asymmetric interactions have on the performance of collaborative manipulations. For that, we compared two *Symmetric* setups, where the interaction and visualization interfaces are the same for both participants, with the *Asymmetric* setup, where interaction and visualization are distinct for each participant. We use the collaborative AR interface described in Sect. 3.1 for the *AR-AR Symmetric* setup. In the *VR-VR Symmetric* setup both participants use the VR technique presented in Sect. 3.2. In the *VR-AR Asymmetric* setup, one participant uses the AR technique while the other uses the VR technique.

In this experiment, we hypothesize that the *VR-VR* condition will be the fastest, the *AR-AR* will be the slowest and the *VR-AR* will be in between the two other conditions. We expect this behavior due to the more intuitive interaction mapping that the VR controllers provide. Regarding social interaction, we expect that in the asymmetric condition, even if participants experience the VE in different realities, they can be aware of each other, communicate and create strategies that take advantage of each interface.

The experiment followed a within-subjects design, with the *Technique Symmetry (VR-VR, AR-AR and VR-AR)* and *Co-manipulation* (manipulating the *Same object* or *Different objects)* as independent variables. We collected the *time* to complete each trial as a dependent variable. The *Manipulation time* of each participant in a dyad is also used to compute a shared *Participation* score dependent variable, which is detailed later in the paper. For the *Technique Symmetry* independent variable, users also answer a SEQ to assess the subjective difficulty of the task and questions of the *Networked Minds Measure of Social Presence* [4] questionnaire about behavioral interaction, mutual assistance and dependent actions.

We analyzed whether the independent variables (*Technique Symmetry* and *Co-manipulation*) and their interaction had a significant effect in the time related dependent variables with a 2-way repeated measures ANOVA test. We verified if the ANOVA assumption of normality of the residuals was violated with the Shapiro-Wilk test. We conducted a Post-hoc analysis with multiple pairwise t-tests of the different variables if a variable with more than 2 levels was found statistically significant. The alpha significance level was set to 0.05.

Lastly, we analyze the effect that the collaborative *VR-VR* condition has on performance and workload when compared with the single user manipulations presented in Sect. 5. To achieve this, we collect the *VR-VR* workload using the NASA TLX questionnaire and interpret data collected across experiments as a between-subject design, with *Group Size* (*Distant* as *Individual* condition and *VR-VR* as *Pairs* condition) as the independent variable. We expect that the performance disadvantage of the *Distant* approach, when compared with the *Virtual Hand*, can be compensated with collaboration. If the workload is reduced in collaboration, this can indicate that the work division plays a role in the collaborative tasks.

6.1 Participants

Thirty-six participants took part in the experiment (11 female), mean age if 21.6 years (SD = 2.19). We asked participants to choose their partner beforehand, since the experiment was conducted in pairs. Three pairs did not know each other beforehand. All subjects read and signed to an informed consent form before the experiment. The majority of participants had none or very little experience with VR headsets, only two of them reported high experience. On the

---

[3] https://humansystems.arc.nasa.gov/groups/tlx/downloads/TLXScale.pdf

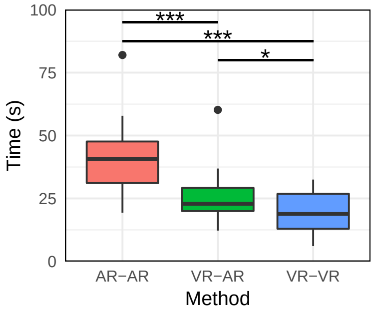


Figure 7: Box and whiskers plot of the trial completion time for each level of the Technique Symmetry variable. *VR-VR* condition is the fastest ($M = 20.93$, $SD = 11.67$) followed by *VR-AR* ($M = 26.62$, $SD = 11.48$) and *AR-AR* ($M = 42.31$, $SD = 16.7$).

other hand, five participants reported low experience with handheld tablets. All participants had either normal or corrected to normal vision. Individuals with corrected vision were instructed to wear their glasses during the VR condition.

6.2 Procedure

Latin square determined the presentation order of the three *Technique Symmetry* conditions (*VR-VR, AR-AR and VR-AR*), resulting in 6 different orders. We explained the operation of the interface at the beginning of each condition. Subjects practiced the manipulations and performed three training trials. Then, they had to complete eight docking trials in sequence. The trials' presentation order was randomized, with four trials in which participants had to manipulate the *Same object* and four trials in which they had to manipulate *Different objects* (*Co-manipulation* variable). We asked the pairs to complete each trial as fast as possible but avoiding to rush, which could lead to inaccuracy. After each *Technique Symmetry* condition, while resting, the subjects answered a single easy question (SEQ) to assess the task difficulty and questions of the *Networked Minds Measure of Social Presence* [4] questionnaire. While in the *VR-VR* condition, participants answered the NASA TLX questionnaire, these answers are used to compare the collaborative *VR-VR* condition with the individual one from the previous experiment (Sect. 5). The experiment took on average 60 minutes. In summary, the experiment design had: **18** pairs × **3** technique symmety conditions × **8** trials (4 - manipulating the same object and 4 - manipulating different objects) = **432** unique trials.

6.3 Results

6.3.1 Trial Completion Time

The statistical analysis indicate that the effect of *Technique Symmetry* was statistically significant ($F_{2,34} = 21.02$, $p < .001$), while it failed to reject equality of the levels of *Co-manipulation* ($F_{1,17} = 2.94$, $p = .157$) and the interaction between the two variables ($F_{2,34} = 1.12$, $p = .339$). Post-hoc analysis comparing the levels of Technique Symmetry suggests that the groups performed the task significantly faster in the *VR-VR* condition as compared to the *AR-AR* ($p < .001$) and *VR-AR* ($p = .018$) conditions. In addition, groups also performed faster in the *VR-AR* condition than in the *AR-AR* condition ($p < .002$). Trial completion time results are presented in Fig. 7.

6.3.2 Error vs. Time Trade-off

Fig. 8 shows the expected time by error reduction for position and rotation of the object and its docking counterpart. The smaller amount of time needed to reduce the error suggests that *VR-VR* is more efficient for position and rotation. We note that the *VR-AR* (asymmetric) condition presented similar performance to *VR-VR*.

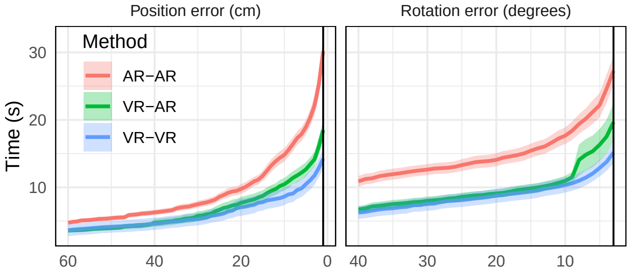


Figure 8: Time taken to reduce the position and rotation errors to the task threshold. The faster the error is reduced, the better. The shaded areas represent the standard error of the mean.

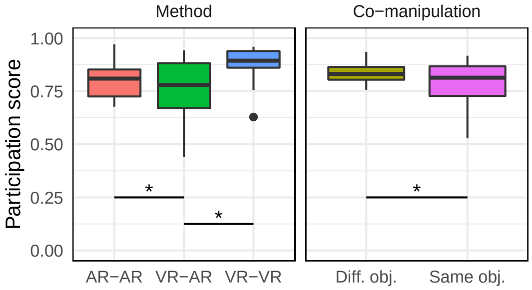


Figure 9: Box and whiskers plot of the participation score according to the Technique Symmetry and Co-manipulation independent variables. There is a significant main effect of both factors.

6.3.3 Participation

We estimate a shared *Participation* score of dyad participants based on the equivalence of manipulation time of each participant. The *Participation* score is computed by:

$$Participation = 1 - abs\left(\frac{Time_{user1} - Time_{user2}}{Time_{user1} + Time_{user2}}\right),$$

where a score of 1 represents an equal time of manipulation, while a score of 0 means that a single participant carried the task for a given piece.

Statistical analysis of the *Participation* score shows that *Co-manipulation* had a significant effect ($F_{1,17} = 6.38$, $p = .022$), with improved participation time balance when users had to control different objects. Moreover, the effect of *Technique Symmetry* was also significant ($F_{2,34} = 4.46$, $p = .019$), with users in the *VR-VR* condition demonstrating improved participation time balance as compared to *VR-AR* ($p = .011$) and *AR-AR* ($p = .023$). Finally, no statistically significant difference was found for the interaction between the two variables ($F_{2,34} = .8$, $p = .457$). *Participation* results are presented in Fig. 9.

6.3.4 Social Presence

We evaluated the social presence with the *Networked Minds Measure of Social Presence* [4]. Non-parametric Friedman tests were conducted to compare the effect of the three *Technique Symmetry* conditions (*VR-VR*, *AR-AR* and *VR-AR*) on three factors of the behavioral engagement dimension of Social Presence, namely: *behavioral interaction*, *dependent actions* and *mutual assistance*. The test failed to reject equality for the all three factors with *behavioral*

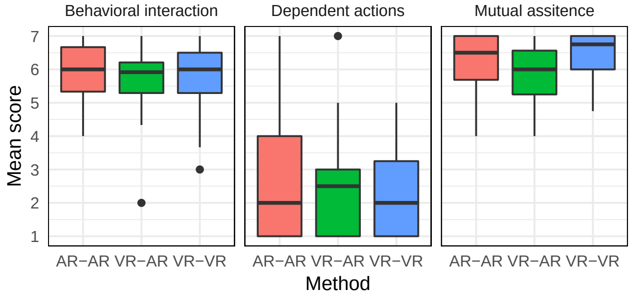


Figure 10: Social presence evaluation. There is no statistically significant difference between conditions for *behavioral interaction*, *dependent actions* and *mutual assistance*.

interaction $p = .179$ ($X^2(2) = 3.48$), *dependent actions* $p = .068$ ($X^2(2) = 5.38$), and *mutual assistance* $p = .051$ ($X^2(2) = 5.95$).

6.3.5 Task Difficulty

We compare the perceived task difficulty per condition with the SEQ "How easy was the task overall?". Non-parametric Friedman was used, with Wilcoxon signed-rank test for pairwise comparisons and Holm-Bonferroni correction for multiple comparisons. A significant difference was found between the levels of *Technique symmetry* ($X^2(2) = 16.43$, $p < .001$), with *VR-VR* considered easier than *AR-AR* ($p = .007$). No statistically significant difference was found between *VR-VR* and *VR-AR* ($p = .073$), or between *VR-AR* and *AR-AR* ($p = .073$).

6.3.6 Individual vs Collaborative Performance and Workload

Fig. 11 shows the trial completion time for users using the *Distant* technique in the first experiment (*individual*) and in the second experiment (*pairs*). Users interacting collaboratively took significantly more time to complete the trials ($t_{28.4} = 2.97$, $p = .006$). On the other hand, the NASA TLX questionnaire indicates that users felt a smaller workload when interacting in pairs ($p = .012$).

7 DISCUSSION

7.1 VR Manipulations

We reported the results of an experiment that compared the *Virtual Hand* with our *Distant* approach. The *Virtual Hand* is considered the most natural interaction paradigm [6], but it has limitations that restrict its use for collaborative activities. As hypothesized, it is faster to dock virtual objects with the *Virtual Hand* approach. Users reported that with more training the *Distant* could outperform the *Virtual Hand* both in efficiency and usability for everyday tasks similarly to what Schultheis et al. [34] concluded. The SUS score reported reinforces this statement.

The results showed that working in pairs did not lead to an increase in speed during the task resolution. We believe that the simplicity of the task may have influenced the results. As we have observed here and in previous works [12, 13], people tend to divide the tasks during collaborative activities. We believe that in tasks with more control variables, with objects having more complex shapes, or that demand carrying the objects over farther distances, the collaborative work can be more helpful in improving performance. After, we investigated whether the participant's workload is affected by working in pairs. As hypothesized, the results indicate that working in pairs significantly reduces the workload. One evidence for this behavior is the result of the participation score reported in the asymmetric experiment. The experiment reveals that both team members constantly participate on solving the manipulation task.

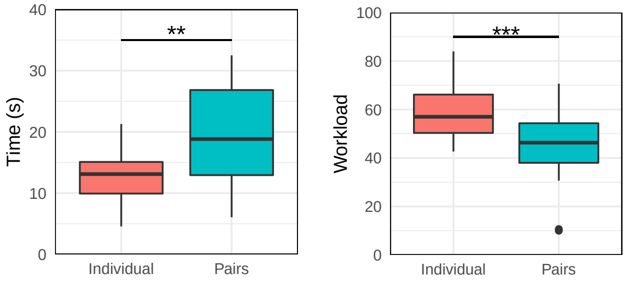


Figure 11: Box and whiskers plot of the trial completion time and workload (NASA TLX) for users using the Distant technique in the first experiment (individual) and in the second experiment (collaborative). (a) *Individuals* completed the tasks faster ($M = 12.69$, $SD = 4.32$) than *Pairs* ($M = 19.38$, $SD = 8.19$), but also (b) reported higher workload ($Mdn = 57$) than *Pairs* ($Mdn = 46.33$).

7.2 Asymmetric Interactions

As hypothesized, we observed that the *VR-VR* condition outperforms both the *AR-AR* and *VR-AR* conditions. The conditions' performance followed a consistent behavior with the asymmetric *VR-AR* being the second fastest and *AR-AR* in the third position. In this analysis, we have evaluated the contribution behavior and coordination behavior listed by [21] as the important collaborative behaviors that affect the team performance. The communication behavior was not evaluated. We investigated the user's contribution during the docking tasks to understand if, while in the asymmetric condition, users tend to participate less when they are in interaction disadvantage, such as using the AR interface. The results showed that regardless of the condition, pairs had similar work participation indicating a high cooperation level even when there is a visualization and interaction asymmetry between the participants.

In the social presence analysis, we observed a slightly lower mutual assistance in the asymmetric condition. The result suggests that the asymmetric condition led to a lower sense that the pairs were helping each other. The reported dominance behavior of the HMD users in the experiment could also be associated with the Field-of-View (FoV) disparity between the VR user and the AR user, as observed by Piumsomboon et al. [26]. While the Oculus Rift used in our experiments has 110° FoV, the iPad's camera has 69.9° diagonal FoV. Even though the users reported lower mutual assistance in the asymmetric setup, we observed that the participants frequently talk about their actions and strategies during the experiment. Such behavior was reflected in the post-experiment questionnaires, where several users reported a work division strategy.

8 CONCLUSION

We explored co-manipulation of virtual objects in asymmetric control and display environments. First, we designed a user interface capable of dealing with simultaneous manipulations in the same virtual object. The technique was developed for HMDs headsets and their controllers. Our approach overcomes the *Virtual Hand* egocentric limitation at the cost of slightly higher manipulation completion time. The advantages of our design become evident in collaborative work. Even though the performance in pairs was lower than that of individuals, the significantly lower workload indicates that users collaborate in a way that the amount of work is distributed among participants, similarly to the findings of the study present in [13]. Second, we explored an asymmetric co-manipulation in a shared environment. VR and AR users coexist in the same physical space while keeping their view perspective, scale and dimensionality. Along with the approach design, we carried out a user study to compare asymmetric and symmetric interactions in a collaborative VE. We vary the reality users experience to assess performance, participation and social presence. The experiment was conducted in pairs in two symmetric conditions, *VR-VR* and *AR-AR* and an asymmetric condition *VR-AR*. We demonstrated that pairs in the *VR-AR* asymmetric condition achieved better performance than the *AR-AR* symmetric condition and slightly worse performance than *VR-VR* symmetric condition. Participants had high work participation for all three conditions, although participation and mutual assistance in the *VR-AR* were significantly lower than in the symmetric conditions. This is evidence that pairs cooperate to solve the task independent of the interaction and visualization interfaces.

8.1 Limitations and Future Work

We limited our augmented reality setup with the use of handheld devices. See-through HMDs, such as the *Hololens* are a promising technology for Mixed Reality (MR). New opportunity of research, with the popularity of such MR interfaces, could include the exploration of the collaborative manipulations in MR scenarios. For that, new interaction design for precise 3D manipulations is required since the current hand gestures interactions are somewhat limited. The exploration of asymmetric setups with HMD VR and See-through HMD such as in the work of Piumsomboon et al. [27] could continue in future works.

We narrowed our study only for object manipulations in 3D virtual environments. Selection is another of the canonical actions in virtual environments and can be further explored during collaborative interactions.

We assessed collaborative interaction with pairs of users, but our solutions support input from several users at the same time. In future works, the collaboration in scenarios with larger groups could be approached in more detail since it is still little explored.

Regarding asymmetric interactions, we explored the interface asymmetry dimension. One promising research venue is the exploration of the scale asymmetry, where the size of the users in the VE changes. The scales can be combined with the interfaces described in this paper. The participants could arrange themselves in different organizational levels depending on the immersion of the technique's visualization screen. At the operational level, where the users are the workers, a portable display or an HMD would support the viewing of individual plans, while at the tactical level, a large display could afford the awareness of the complete task state and formulation of a strategy that includes multiple workers [2, 8]. Further studies could also approach if the relationship of participants influences the group performance and the strategic decisions.

ACKNOWLEDGMENTS

We wish to thank the anonymous reviewers for their comments and all the users that volunteered for the experiment. Thanks are also due to Aimée Calepso for her participation in the video recordings. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - FinanceCode 001 and in part by CNPq-Brazil. We also acknowledge FAPERGS (project 17/2551-0001192-9) and CNPq-Brazil (project 311353/2017-7) for their financial support.

REFERENCES

[1] L. Aguerreche, T. Duval, and A. Lécuyer. Reconfigurable tangible devices for 3D virtual object manipulation by single or multiple users. In *Proceedings of the 17th ACM Symposium on Virtual Reality Software and Technology*, page 227, New York, USA, 2010. ACM Press.

[2] F. Bacim, E. D. Ragan, C. Stinson, S. Scerbo, and D. A. Bowman. Collaborative navigation in virtual search and rescue. In *2012 IEEE Symposium on 3D User Interfaces (3DUI)*, pages 187–188. IEEE, 2012.

[3] A. Bangor, P. Kortum, and J. Miller. Determining what individual sus scores mean: Adding an adjective rating scale. *J. Usability Studies*, 4(3):114–123, May 2009.

[4] F. Biocca, C. Harms, and J. Gregg. The networked minds measure of social presence: Pilot test of the factor structure and concurrent validity. *4th Annual Int. Workshop on Presence*, pages 1–9, 2001.

[5] D. A. Bowman and L. F. Hodges. An evaluation of techniques for grabbing and manipulating remote objects in immersive virtual environments. *Proceedings of the 1997 symposium on Interactive 3D graphics - SI3D '97*, pages 35–ff., 1997.

[6] E. L. J. J. P. I. Bowman A. D; Kruijff. *3D User Interfaces Theory and Pratice*. Ed Addison-wesley, Boston, MA, 2005.

[7] J. Brooke. SUS: a retrospective. *Journal of usability studies*, 8(2):29–40, 2013.

[8] M. L. Chenechal, J. Lacoche, J. Royan, T. Duval, V. Gouranton, and B. Arnaldi. When the giant meets the ant an asymmetric approach for collaborative and concurrent object manipulation in a multi-scale environment. In *2016 IEEE Third VR International Workshop on Collaborative Virtual Environments*, pages 18–22. IEEE, 2016.

[9] E. F. Churchill and D. Snowdon. Collaborative virtual environments: an introductory review of issues and systems. *Virtual Reality*, 3(1):3–15, 1998.

[10] T. Duval, A. Lécuyer, and S. Thomas. SkeweR: A 3D interaction technique for 2-user collaborative manipulation of objects in virtual environments. In *3DUI 2006: IEEE Symposium on 3D User Interfaces 2006 - Proceedings*, volume 2006, pages 69–72, 2006.

[11] T. Goldschwendt, C. Anthes, G. Schubert, D. Kranzlmüller, and F. Petzold. The collaborative design platform protocol. In *Proceedings of the 20th ACM Symposium on Virtual Reality Software and Technology - VRST '14*, pages 239–240, New York, NY, USA, 2014. ACM Press.

[12] J. G. Grandi, H. G. Debarba, L. Nedel, and A. Maciel. Design and Evaluation of a Handheld-based 3D User Interface for Collaborative Object Manipulation. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI '17*, pages 5881–5891, New York, New York, USA, 2017. ACM Press.

[13] J. G. Grandi, H. G. Debarba, L. Nedel, and A. Maciel. Design and Assessment of a Collaborative 3D Interaction Technique for Handheld Augmented Reality. In *2018 IEEE Virtual Reality (VR)*. IEEE, 2018.

[14] J. Gugenheimer, E. Stemasov, J. Frommel, and E. Rukzio. ShareVR: Enabling Co-Located Experiences for Virtual Reality between HMD and Non-HMD Users. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pages 4021–4033, 2017.

[15] H. Ibayashi, Y. Sugiura, D. Sakamoto, N. Miyata, M. Tada, T. Okuma, T. Kurata, M. Mochimaru, and T. Igarashi. Dollhouse VR: A multi-view, multi-user collaborative design workspace with VR technology. *SIGGRAPH Asia 2015 Posters, SA 2015*, pages 2–3, 2015.

[16] J. J. LaViola Jr, E. Kruijff, R. P. McMahan, D. Bowman, and I. P. Poupyrev. *3D user interfaces: theory and practice*. Addison-Wesley Professional, 2017.

[17] G. A. Lee, T. Teo, S. Kim, and M. Billinghurst. A user study on mr remote collaboration using live 360 video. In *2018 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 153–164. IEEE, 2018.

[18] D. P. Mapes and J. M. Moshell. A two-handed interface for object manipulation in virtual environments. *Presence: Teleoper. Virtual Environ.*, 4(4):403–416, Jan. 1995.

[19] D. Margery, B. Arnaldi, and N. Plouzeau. A general framework for cooperative manipulation in virtual environments. *Virtual Environments*, 99:169–178, 1999.

[20] A. Marzo, B. Bossavit, and M. Hachet. Combining multi-touch input and device movement for 3D manipulations in mobile augmented reality environments. In *Proceedings of the Symposium on Spatial User Interaction*, pages 13–16, New York, USA, 2014. ACM Press.

[21] M. M. Montoya, A. P. Massey, and N. S. Lockwood. 3D Collaborative Virtual Environments: Exploring the Link between Collaborative Behaviors and Team Performance. *Decision Sciences*, 42(2):451–476, 2011.

[22] A. Mossel, B. Venditti, and H. Kaufmann. 3DTouch and HOMER-S: Intuitive Manipulation Techniques for One-Handed Handheld Augmented Reality. *Proceedings of the Virtual Reality International Conference on Laval Virtual - VRIC '13*, page 1, 2013.

[23] C. Nguyen, S. DiVerdi, A. Hertzmann, and F. Liu. CollaVR. *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology - UIST '17*, pages 267–277, 2017.

[24] T. T. H. Nguyen, T. Duval, and C. Fleury. Guiding Techniques for Collaborative Exploration in Multi-Scale Shared Virtual Environments. *GRAPP International Conference on Computer Graphics Theory and Applications*, pages 327–336, 2013.

[25] M. S. Pinho, D. A. Bowman, and C. M. D. S. Freitas. Cooperative object manipulation in collaborative virtual environments. *Journal of the Brazilian Computer Society*, 14:54–67, 2008.

[26] T. Piumsomboon, A. Day, B. Ens, Y. Lee, G. Lee, and M. Billinghurst. Exploring enhancements for remote mixed reality collaboration. In *SIGGRAPH Asia 2017 Mobile Graphics & Interactive Applications*, SA '17, pages 16:1–16:5, New York, NY, USA, 2017. ACM.

[27] T. Piumsomboon, G. A. Lee, J. D. Hart, B. Ens, R. W. Lindeman, B. H. Thomas, and M. Billinghurst. Mini-Me. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18*, pages 1–13, New York, New York, USA, 2018. ACM Press.

[28] K. Riege, T. Holtkamper, G. Wesche, and B. Frohlich. The bent pick ray: An extended pointing technique for multi-user interaction. *3DUI 2006: IEEE Symposium on 3D User Interfaces 2006 - Proceedings*, 2006:62–65, 2006.

[29] M. Robinson. Designing for Unanticipated Use... In *ECSCW '93*, pages 187–202, 1993.

[30] J. S. Roo and M. Hachet. One Reality: Augmenting How the Physical World is Experienced by combining Multiple Mixed Reality Modalities. *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology - UIST '17*, pages 787–795, 2017.

[31] R. A. Ruddle, J. C. D. Savage, and D. M. Jones. Symmetric and asymmetric action integration during cooperative object manipulation in virtual environments. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 9(4):285–308, 2002.

[32] H. Salzmann, J. Jacobs, and B. Froehlich. Collaborative interaction in co-located two-user scenarios. In *Proceedings of the 15th Joint Virtual Reality Eurographics Conference on Virtual Environments*, JVRC'09, pages 85–92, Aire-la-Ville, Switzerland, Switzerland, 2009. Eurographics Association.

[33] J. Sauro and J. S. Dumas. Comparison of three one-question, post-task usability questionnaires. In *Proceedings of the Conference on Human Factors in Computing Systems*, pages 1599–1608. ACM, 2009.

[34] U. Schultheis, J. Jerald, F. Toledo, A. Yoganandan, and P. Mlyniec. Comparison of a two-handed interface to a wand interface and a mouse interface for fundamental 3d tasks. In *3D User Interfaces (3DUI), 2012 IEEE Symposium on*, pages 117–124. IEEE, 2012.

[35] R. N. Shepard and J. Metzler. Mental Rotation of Three-Dimensional Objects. *Science*, 171(3972):701–703, feb 1971.

[36] E. Sundén, I. Lundgren, and A. Ynnerman. Hybrid Virtual Reality Touch Table - An Immersive Collaborative Platform for Public Explanatory Use of Cultural Objects and Sites. pages 109–113, 2017.

[37] A. C. Ulinski, Z. Wartell, P. Goolkasian, E. A. Suma, and L. F. Hodges. Selection performance based on classes of bimanual actions. *3DUI - IEEE Symposium on 3D User Interfaces 2009 - Proceedings*, pages 51–58, 2009.

[38] J. Wang, R. Budhiraja, O. Leach, R. Clifford, and D. Matsuda. Escape from Meadwyn 4: A cross-platform environment for collaborative navigation tasks. *IEEE Symposium on 3D User Interfaces 2012, 3DUI 2012 - Proceedings*, pages 179–180, 2012.

[39] A. Zenner, F. Kosmalla, M. Speicher, F. Daiber, and A. Krüger. A projection-based interface to involve semi-immersed users in substitutional realities. In *2018 IEEE 4th Workshop on Everyday Virtual Reality (WEVR). Workshop on Everyday Virtual Reality (WEVR-2018), 4th, March 18, Reutlingen, Germany*. IEEE, 3 2018.