Virtual Grasping Feedback and Virtual Hand Ownership

Ryan Canales*
Clemson University
rcanale@clemson.edu

Aline Normoyle
Venturi Labs & Swarthmore College
aline@venturilabs.com

Yu Sun
Clemson University
ysun3@clemson.edu

Yuting Ye
Facebook Reality Labs
yuting.ye@oculus.com

Massimiliano Di Luca
Facebook Reality Labs & University of Birmingham UK
max.diluca@oculus.com

Sophie Jörg
Clemson University
sjoerg@clemson.edu

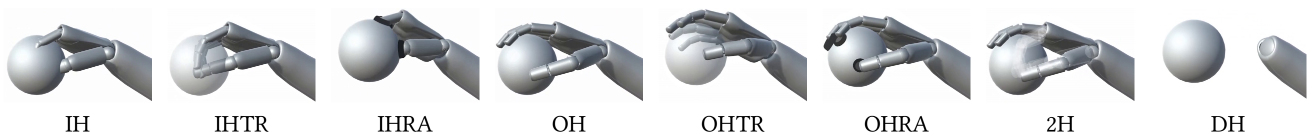


Figure 1: Grasping feedback techniques used in this study: Inner Hand (IH, displays the tracked hand), Inner Hand with Object Transparency (IHTR, object becomes transparent when grasped), Inner Hand with Reactive Affordance (IHRA, visualizations aiming at a more tactile feeling are added), Outer Hand (OH, virtual hand does not penetrate the object), Outer Hand with Object Transparency (OHTR) and Outer Hand with Reactive Affordance (OHRA), Two Hands (2H, visualizes the tracked hand and the outer hand), and Disappearing Hand (DH, virtual hand disappears during the grasp).

ABSTRACT

In this work, we investigate the influence of different visualizations on a manipulation task in virtual reality (VR). Without the haptic feedback of the real world, grasping in VR might result in intersections with virtual objects. As people are highly sensitive when it comes to perceiving collisions, it might look more appealing to avoid intersections and visualize non-colliding hand motions. However, correcting the position of the hand or fingers results in a visual-proprioceptive discrepancy and must be used with caution. Furthermore, the lack of haptic feedback in the virtual world might result in slower actions as a user might not know exactly when a grasp has occurred. This reduced performance could be remediated with adequate visual feedback.

In this study, we analyze the performance, level of ownership, and user preference of eight different visual feedback techniques for virtual grasping. Three techniques show the tracked hand (with or without grasping feedback), even if it intersects with the grasped object. Another three techniques display a hand without intersections with the object, called outer hand, simulating the look of a real world interaction. One visualization is a compromise between the two groups, showing both a primary outer hand and a secondary tracked hand. Finally, in the last visualization the hand disappears during the grasping activity.

In an experiment, users perform a pick-and-place task for each feedback technique. We use high fidelity marker-based hand tracking to control the virtual hands in real time. We found that the tracked hand visualizations result in better performance, however, the outer hand visualizations were preferred. We also find indications that ownership is higher with the outer hand visualizations.

CCS CONCEPTS

• **Human-centered computing → Virtual reality**; **Visualization**.

KEYWORDS

virtual grasping, visual feedback, virtual hand interaction, virtual reality, body ownership, virtual character

ACM Reference Format:
Ryan Canales, Aline Normoyle, Yu Sun, Yuting Ye, Massimiliano Di Luca, and Sophie Jörg. 2019. Virtual Grasping Feedback and Virtual Hand Ownership. In *ACM Symposium on Applied Perception 2019 (SAP '19), September 19–20, 2019, Barcelona, Spain.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3343036.3343132

1 INTRODUCTION

Object manipulation in virtual environments plays a significant role in the overall user experience and many methods have been suggested to interact with objects in immersive virtual reality (IVR) [Bowman and Hodges 1997; Poupyrev et al. 1997]. In IVR applications, the degree to which real world interactions are reproduced [McMahan et al. 2012], is a goal that many researchers and designers aim for. Though natural interaction techniques may not have the highest accuracy or performance, they can increase intuitiveness and immersion [Bowman et al. 2012; Lin et al. 2019]. Therefore,

* This is the corresponding author

SAP '19, September 19–20, 2019, Barcelona, Spain

ACM ISBN 978-1-4503-6890-2/19/09…$15.00
https://doi.org/10.1145/3343036.3343132

many applications use a virtual hand metaphor for interaction as it is closest to how we interact with objects in the real world. As hand tracking hardware continues to advance, it becomes increasingly important to study visual feedback techniques for making it easier and more intuitive to interact with virtual environments using one's hands. To this aim, some methods use hand tracking and simulate interactions using physically based methods for grasping [Borst and Indugula 2005; Zhao et al. 2013], sometimes in combination with data driven approaches [Liu 2008; Pollard and Zordan 2005].

Despite all simulations, our real hands and fingers are typically not prevented from going through virtual objects. Rendering the accurate position of our hands and fingers in such cases creates finger object interpenetration in the virtual world, which reduces realism and can degrade the immersive experience [Prachyabrued and Borst 2012]. When manipulating objects in IVR, it is possible to render a hand that remains outside of the geometry, a representation that is closer to our real world experience. However, as the hand or fingers would not be displayed in accordance with the users' accurate motions, such a technique might result in visual-proprioceptive discrepancy [Prachyabrued and Borst 2013] and a reduced sense of control of the motions of the virtual character, which might reduce the feeling of embodiment.

In this study, we investigate which option - an outer hand that does not intersect with virtual objects or an inner hand which accurately follows the users' motions - is more advantageous for the user regarding performance, ownership, and preference. Furthermore, in IVR the user has no tactile feedback to know when exactly a grasping action has started successfully. As additional grasping feedback can increase performance [Prachyabrued and Borst 2012, 2014], we add examples with secondary feedback into our experimental design. Finally, we wonder if it is possible to increase the feeling of ownership using grasping visualizations and designed a condition with this aim.

In an experiment, we investigate the effect of eight different visual feedback techniques for virtual grasping (shown in Figure 1, on performance, perceived ownership, and user preference. Participants are seated in front of a table in a virtual environment and perform a simple pick-and-place task. Our study builds upon previous work by not only measuring task performance, but also examining the influence that the visualizations have on virtual hand embodiment. Furthermore, compared to previous work, our novel hand tracking system allows for real-time, highly precise hand tracking in a relatively large space.

2 RELATED WORK

In this section, we present related work on the effect of visual feedback for manipulation tasks on performance. In our study, we also evaluate if the selected visualizations affect ownership as embodiment is often a goal of controlled avatars in virtual environments. While previous research has not studied the effect of visualization on embodiment specifically, we present several studies which have examined the effect of hand appearance on ownership.

2.1 Visual Feedback for Grasping

Appropriate visual feedback for interaction in virtual environments can not only enhance the user experience but it can also affect efficiency [Argelaguet and Andújar 2013]. Lam et al. [2018] tested virtual grasping in a desktop environment and found that a grasping animation as visual feedback helped participants notice when an object is selected. Vosinakis and Koutsabasis [2018] tested the grasp and release performance of multiple visual feedback techniques in a desktop environment and in IVR. They found that bare hand grasping is performed best in IVR and that any form of visual feedback resulted in better grasping and release performance than none. Geiger et al. [2018] tested visual feedback for assisting users in gripping a virtual object in a specified way. Their results showed that for complex grip types, such as whole hand grasping, visual feedback significantly improved user performance.

Prachyabrued and Borst's work [2014] in evaluating visual feedback for grasping is closest to ours. They evaluated the performance and subjective preference of eight different visual cues for finger interpenetration during manipulation. The techniques tested were called *Inner Hand* (IH), *Outer Hand* (OH), *See Through* (ST), *2-Hand* (2H), *Finger Color* (FC), *Object Color* (OC), *Arrow* (AR), and *Vibration* (VB). IH (tracked hand), OH (virtual fingers stay outside of the object), and 2H (IH and OH combined) are similar to the conditions with the same names in our study. The FC, OC, AR, and VB techniques each give indirect feedback for interpenetration. The color of the fingers or object changes based on depth for FC and OC, respectively. In AR, arrows extend from the points of contact and change in length as a function of depth. In VB, the virtual fingers vibrate as the tracked hand enters the virtual object. Participants used each technique to grasp a virtual ball and release it over a target. IH was found to be the best for performing the ball drop accurately in contrast to OH, which was the worst. 2H was notably a good compromise, as it generally resulted in better performance than the others. Visual appearance had a significant impact on the users' preference, with OC and FC being the most preferred, followed by OH, 2H, and AR, then ST and IH, and finally VB.

Though it is well established that good visual feedback for virtual grasping is helpful in terms of interaction performance and user preference, it is important to consider the visual feedback from a presence and immersion standpoint. In contrast to Prachyabrued and Borst [2014], we measure not only the performance for different visual feedback techniques, but also the effect that they have on ownership of the virtual hand.

2.2 Virtual Hand Embodiment

Virtual reality has allowed researchers to study the extent to which we can establish a sense of embodiment over virtual avatars. Kilteni et al. [2012a] define the sense of embodiment (SoE) as follows: *SoE toward a body B is the sense that emerges when B's properties are processed as if they were the properties of one's own biological body.* There are three components that contribute to SoE: body ownership, self-location, and the sense of agency [Kilteni et al. 2012a; Longo et al. 2008]. Ownership refers to the feeling that the virtual body is one's own body. Location is the feeling that one's body and the virtual body are in the same place. Agency is the feeling that one has control over the virtual body.

Slater et al. [2008] induced the *Virtual Arm Illusion*, a sense of embodiment towards a virtual arm, using tactile stimulation on the real hand and both synchronous and asynchronous virtual visual stimuli. The results were that synchronized visual and tactile stimuli result in significantly higher levels of ownership. Ma and Hommel [2013] subject the virtual hand to a threat, in which the virtual hand is cut with a knife, and an impact, where a ball hits the virtual hand. Though ownership was higher when the user had synchronous control over the virtual hand, asynchronous control (the virtual hand movement was delayed) had no effect on the users' emotional investment in the threat condition.

Since accurate user control of the virtual hand has been correlated with ownership, researchers also began investigating the relationship between the appearance of the hand and ownership. Many different visual appearances have been investigated from realistic arms and an abstract arrow [Yuan and Steed 2010], over an arm with three times the length of a real arm [Kilteni et al. 2012b] and a rectangle [Ma and Hommel 2015a], to toony, zombie, or robot arms, or even a wooden block [Lin and Jörg 2016]. The general trend is that while some level of ownership occurs for all representations, ownership is typically stronger with anthropomorphic representations and strongest with a realistic representation. Argelaguet et al. [2016] conducted a study to investigate the effect that virtual hand representation has on users' sense of agency and sense of ownership of the hand. They used three visual representations: an abstract sphere, an iconic hand composed of simple shapes, and a highly realistic hand. The users were asked to move a ball over virtual hazards, including a flame, a spinning saw, and barbed wire. The results were that the realistic hand elicited the strongest sense of ownership. For performance and agency, however, the abstract representations were superior. Though previous studies have demonstrated that agency and ownership are correlated [Ma and Hommel 2015b], the reduced sense of agency over the realistic hand was not strong enough to affect ownership.

In our experiment we expect that our high fidelity, marker-based real time hand tracking system will lead to high agency and therefore high ownership levels overall.

3 METHOD

For this study, we created a simple virtual test room containing a desk, a chair, and a humanoid robot avatar using the Unity game engine. Participants are seated in front of a real table and are represented in VR as a robot avatar. A virtual button is centered on the desk in front of the avatar with a virtual ball on one side of it, and a target (a red X) on the other side, as shown in Figure 2 (left). The participants perform a simple pick-and-place task in which they pick up the ball and move it to the target several times with each grasping feedback technique. The black sign on the virtual desk in Figure 2 (left) glows after the start button is pressed to signal to the user that they may begin the task, and likewise dims when the user has finished the task. While performing the task, two different virtual threats (Figure 2, right) occur at separate times.

3.1 Design

The experiment is conducted within subjects, with each participant performing the pick and place task with their dominant hand. The visual feedback conditions are presented in a different random order for each participant.

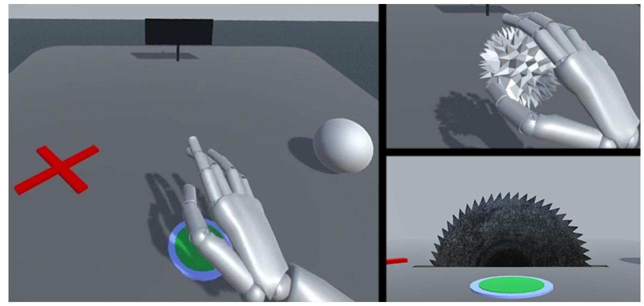


Figure 2: Virtual desk setup for the pick and place task (left) with the virtual threats (right). The spiky ball (top) is displayed twice at random during the second block. The spinning saw (bottom) is displayed in the third block.

3.2 Visualizations

We investigate eight different visualizations for virtual grasping, which we call *Inner Hand* (IH), *Inner Hand with Object Transparency* (IHTR), *Inner Hand with Reactive Affordance* (IHRA), *Outer Hand* (OH), *Outer Hand with Object Transparency* (OHTR), *Outer Hand with Reactive Affordance* (OHRA), *Two Hands* (2H), and *Disappearing Hand* (DH), and which are described and grouped as follows:

Tracked Hand Group:

(1) *IH:* The virtual hand is always controlled by the user's tracked hand and can penetrate the virtual ball.
(2) *IHTR:* Displays the Inner Hand and the virtual ball becomes semitransparent upon being grasped and opaque upon release.
(3) *IHRA:* Displays the Inner Hand and a "dimple mesh" is rendered at the projected contact points when the virtual hand is within a certain distance from the ball, and the dimple grows in size as the user tightens their grip.

Outer Hand Group:

(1) *OH:* The virtual hand always remains outside of the virtual ball. That means that when the tracked hand penetrates the ball during a grasp, the displayed hand does not follow the user's motions.
(2) *OHTR:* Same as IHTR, but the virtual hand remains outside of the virtual ball.
(3) *OHRA:* Same as IHRA, but the virtual hand remains outside of the virtual ball.

Other:

(1) *2H:* The primary virtual hand remains outside of the virtual ball (OH) and the tracked hand (IH) is displayed as a secondary virtual hand as the user grasps the ball.
(2) *DH:* The virtual hand disappears once the ball is grasped and reappears when released.

The Inner Hand and Outer Hand visualizations are included to answer our question if either a visual-proprioceptive discrepancy or hand-object interpenetrations would create a better experience for the user when looking at performance, ownership, and preference. IH and OH were also used in Prachyabrued and Borst [2014], where IH was found to have the best release performance while OH was found to have the worst. Still, OH was preferred more than IH.

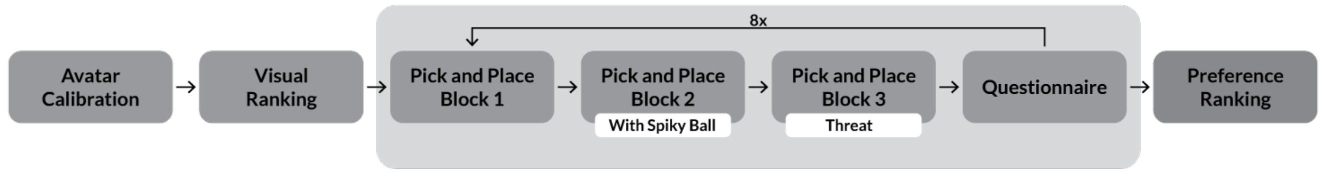


Figure 3: Experimental procedure of our study.

We add the 2H condition to our study as it was considered a good compromise solution in that work.

The visualizations using transparency (IHTR and OHTR) are included to add accurate feedback about when a successful grasp has happened to increase performance. The reactive affordance (IHRA and OHRA) visualizations were added to see if we could increase perceived ownership with visual feedback. It was inspired by Schubert and Fox's work [2017]. We expect that this feedback can help guide the user toward grasping the ball and that it may provide a more tactile feeling. IHRA and OHRA also provide indirect feedback for hand-object interpenetration by changing the size and height of the dimple mesh based on the depth of the fingertip associated with it.

Finally, we include the disappearing hand (DH) visualization as it is a standard method used in many VR games [Owlchemy Labs 2016; Schell Games 2016]. It furthermore keeps the hand from occluding the ball, which could help the user when placing the ball accurately.

3.3 Participants

Twenty-three participants (11F, 12M, ages 18-60, median age group 26-30) with a varied range of previous experience in VR, took part in our experiment. At the beginning of the experiment we obtained signed consent from all participants and pre-screened each participant for cybersickness. After completing the study, participants were debriefed and received a $10 voucher for their time.

3.4 Apparatus

Our virtual reality setup consists of an Oculus Rift CV1 Head-Mounted-Display (HMD), 16 OptiTrack motion capture cameras mounted on support beams surrounding the user on four sides, and a small table in front of the user (Figure 4 (a)). We used real time marker based hand tracking [Han et al. 2018] with a dense marker set of 19 markers per hand. The markers are attached to six gloves of different sizes. The system was pre-calibrated for use with each glove size, so that the avatar's hand size is adjusted to the used glove. The system tracks hand motions at 120FPS, and the VR application runs at 90FPS.

3.5 Grasping Implementation

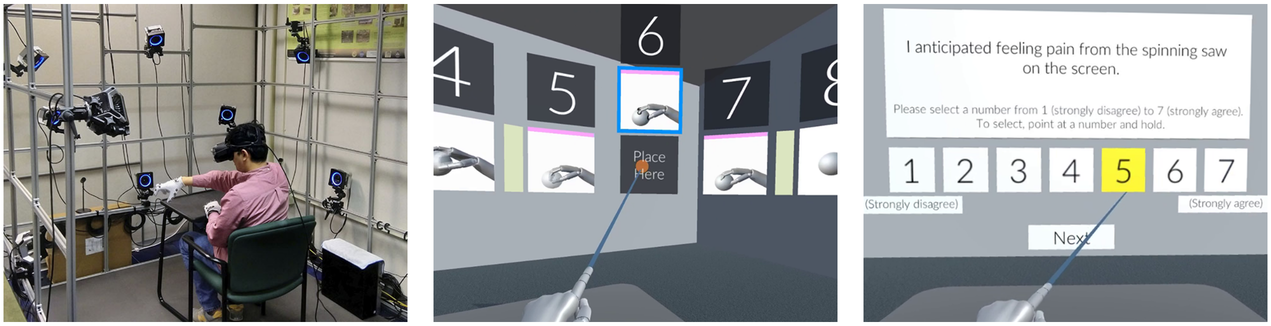
For detecting when the user grasps and releases the ball, we use a heuristic algorithm that was carefully adjusted in a series of tests. A grasp is registered when the tip of the thumb and at least two other fingertips of the tracked hand have contacted the ball. Then, the ball is following the motion of the hand. A release is registered when no fingertips of the tracked hand are touching the ball. We compute two virtual hands; a *visible hand* and a *tracked hand* that always follows the user's movements but is hidden. For all inner hand conditions, the visible hand simply follows the movement of the tracked hand. For the outer hand conditions, the visible hand follows the tracked hand until a grasp has been detected. When a grasp has been detected, the finger joints of the visible hand can rotate or not based on whether or not a joint is contacting the ball and its position in the joint hierarchy for the finger. The joint hierarchy for each finger is: MCP, PIP, DIP, then fingertip, with each joint being the child of the previous. If a joint collides with the ball, then all the joints above it in the hierarchy are locked and the joints below it in the hierarchy can rotate up to 60 degrees towards the ball relative to its parent or until they collide with the ball. This method ensures that the fingers stay outside of the ball's geometry and that a natural grasp pose is reached. To ensure that the hand remains outside of the ball regardless of how quickly the participants grasp, the interpolation speed of the outer hand joints are scaled down as a function of their distance to the surface of the ball. Test subjects were unable to notice a difference in speed.

3.6 Procedure

A visual overview of the experimental procedure is displayed in Figure 3. After agreeing to participate in the study, participants choose the pair of motion tracking gloves that best conform to their hand size and the tracking system and avatar are set up to track and display the corresponding hand size. They are seated in a chair in front of a small desk and fitted with an HMD. Finally, the virtual avatar, sitting in front of a virtual desk, is calibrated such that the arm span and height of the avatar closely match those of the participant.

After calibration, the participants are asked to rank the different grasping visualizations in order of visual appeal based on videos of the interaction (Figure 4 (b)). Once they are finished, a virtual ball appears on the desk and the participants can practice grasping and releasing it with a randomly selected visualization condition. After indicating that they feel comfortable, eight practice pick-and-place trials are started, one with each visualization, in random order. Participants perform the experiment with their dominant hand and complete it in 35 to 60 minutes.

The main experimental task, referred to as the "pick-and-place task", is performed as follows: First, the participant presses a green button on the virtual table, triggering a stopwatch. After pressing the button, the participant can pick up the ball and move it to the target. Once the ball makes contact with the target, the target turns white and the virtual button becomes red. The participant releases the ball on the target and presses the button to stop the timer. After



(a) (b) (c)

Figure 4: (a) Experimental setup with real time marker based hand tracking, a 16 camera optical motion capture system, and an Oculus Rift CV1 HMD. (b) User interface for ranking in VR. The selected visualization (middle with blue outline) and the adjacent ones play a video of the same interaction under their respective visualization condition. The user can also point at one of the other images to play the video for that visualization. (c) Selecting an answer in the VR questionnaire.

two seconds, the scene resets, and the procedure is repeated. The participants were instructed to perform the task as efficiently as they could.

The experiment is divided into eight sessions, one per grasping visualization, and each session is divided into three blocks. During the first block, the pick-and-place task is repeated ten times with the smooth ball. In the second and third block we add dangers or threats to the experiment to measure the reaction of the participant as an indicator of ownership and agency [Argelaguet et al. 2016; Lin and Jörg 2016]. We add two types of dangers, a spinning saw affecting the whole hand as it has been used in previous research [Argelaguet et al. 2016] and a spiky ball that would specifically harm the fingers when grasping. Therefore, the second block resembles the first block, except that the spiky ball replaces the smooth ball for two randomly selected trials out of the ten. In the third block, a spinning saw appears between the ball position and the target position, and the pick and place task is performed once more with the smooth ball.

After the third block, a questionnaire is presented in VR, and participants can select their responses by pointing to the appropriate number (see Figure 4 (c)). Then the trial for the next hand visualization begins. Once the eighth session is complete, the participants rank the visualizations in order of overall preference using the same procedure used for ranking based on appearance.

3.7 Hypotheses

Based on related work, we formed the following hypotheses:

- **H1:** Visualizing the tracked hand will result in better performance than if only the outer hand is displayed.
- **H2:** Feedback of the grasp and release states will improve performance.
- **H3:** Visualizing the tracked hand will result in a stronger sense of embodiment.
- **H4:** The reactive affordance visualization will increase the sense of embodiment.
- **H5:** Visualizations without intersections (outer hand) will be visually preferred.

4 RESULTS

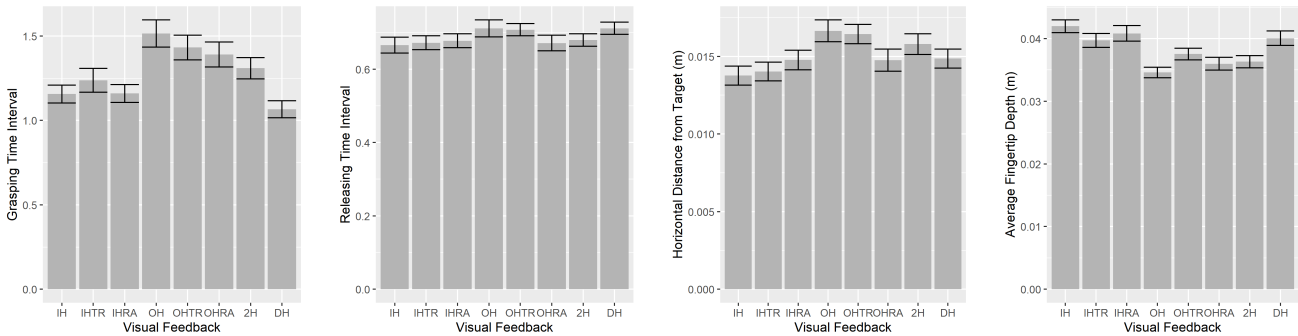
4.1 Performance

Our evaluation of performance is based on grasp performance and release performance. Our sole measure for grasp performance is the grasping time interval, which is the time required to perform a successful grasp. We gauge release performance using the release time interval and the placement accuracy. We furthermore examine the average fingertip depth before release, which is the horizontal distance from the target once the ball is released to give us indications on the participants' behavior and since it has been interpreted as a contributor to release effects [Prachyabrued and Borst 2012, 2014].

Performance is measured in the first block of the pick-and-place task for each visualization condition. The data is first filtered to remove outliers in overall completion time ($\pm 2\sigma$). Ninety-five percent of the data is retained after filtering. For each of the four measurements, a one-way, repeated measures ANOVA was performed, with Greenhouse-Geisser sphericity corrections, with the measurement as the dependent variable and visualization as the independent variable. For post-hoc testing, we use a linear mixed effects model with the performance measure as the dependent variable, visualization as the predictor, and participants and visualization as random effects. The Tukey test was used to make pairwise comparisons of the visualizations on the model.

To test our hypothesis **H1**, for each performance measure we also compared all conditions in the tracked hand group (IH, IHTR, IHRA) to those in the outer hand group (OH, OHTR, and OHRA), again using a one-way repeated measures ANOVA with the group as the independent variable and the performance as the dependent variable.

Grasp Performance: To measure the full action of grasping starting before the grasp until after the user realizes that the grasp was successful, which might include multiple grasp attempts, we define a spherical grasping area ($r = 25$cm) with the ball ($r = 5.48$cm) at its center. We measure the interval between the time when the base of the palm of the virtual hand enters the grasping area until it leaves it.



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

Figure 5: Performance measures results: (a) Grasping time interval includes reaching for the ball and moving it. The time interval for grasping for the outer hand conditions is significantly longer than for the tracked hand conditions. (b) Release time interval, includes reaching the target and moving away from the ball after releasing it. (c) Placement accuracy, horizontal distance of the ball from the target. (d) Average depth of the fingertips before release. Penetration depth was lower for the outer hand group than for the tracked hand group.

We found a significant main effect of visualization on the grasping time interval ($F(7, 154)$, $p < 0.05$). The post-hoc test showed significant differences ($p < 0.05$) between OH/IH, OH/IHRA, OH/DH, and OHTR/DH, with tendencies ($p < 0.1$) toward decreasing time between OH/IHTR and OHRA/DH (Figure 5 (a)). Comparing the tracked hand and outer hand groups showed significantly longer grasp time intervals for the outer hand group ($F(1, 22)$, $p < 0.05$).

Release Performance: We define a spherical release area (r = 25cm) with the ball at its center. Release time interval is computed as the time interval from when the ball contacts the target to when the base of the palm of the virtual hand exits the release area. This interval also captures the time taken for the users to realize that a release happened. Placement accuracy is the horizontal distance between the ball and the target after release. The fingertip depth measures how far into the ball the fingertip has reached, and the average fingertip depth is the mean depth of the five fingertips. We analyze the average fingertip depth at the time the ball contacts the target to capture how tight the user's grip is just before release.

For the release time interval, there were no statistically significant results between the conditions nor between the tracked hand and outer hand groups (Figure 5 (b)).

Though no statistically significant differences were found between all conditions for placement accuracy, the general trend ($p \approx 0.1$) follows our hypothesis that tracked hand conditions tend to outperform outer hand conditions. This trend is shown in Figure 5 (c), with IH resulting in the lowest displacement, and OH resulting in the highest.

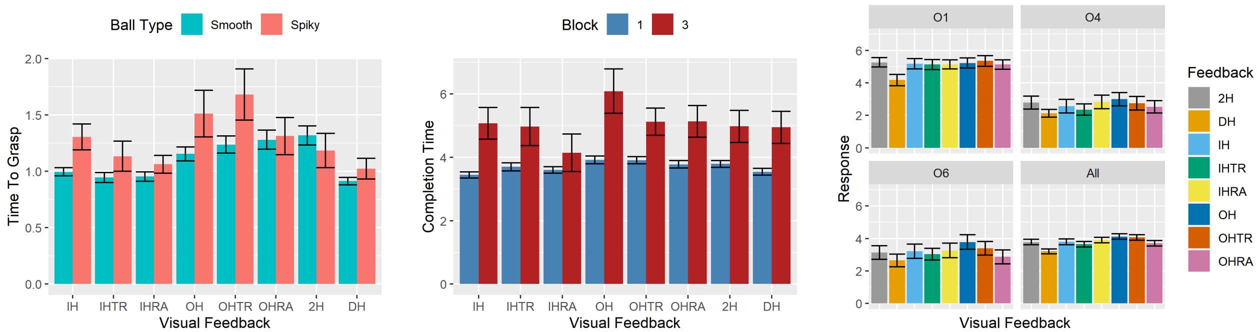
We found a significant effect of visualization for the average fingertip depth before release ($F(7, 154)$, $p < 0.05$, Figure 5 (d)). The post-hoc Tukey test showed significant differences ($p < 0.05$) between OH/IH, OHRA/IH, OH/IHRA and DH/OH with tendencies ($p < 0.1$) for OH/IHTR. A main effect was furthermore found between the tracked and outer hand visualizations, with the outer hand visualizations resulting in a looser grip ($F(1, 22)$, $p < 0.05$).

Summary: Overall, the results for grasping and release performance **confirm our hypothesis H1**, that visualizing the tracked hand will result in better performance. None of the post-hocs showed significant differences between the reactive affordance conditions and their base conditions (IHRA and IH or OHRA and OH) or the transparency feedback conditions and their base conditions (IHTR and IH or OHTR and OH). Therefore, we **can not confirm H2**, that feedback of the grasp or release states will improve performance. Interestingly, 2H did not differ significantly from any of the others for any of our performance measures and its performance was mostly between the performances of the inner hand and outer hand visualizations.

4.2 Ownership

To assess the effect of our visualizations on perceived ownership, we evaluate the impact of the two threats (the spiky ball and the saw) as well as participants' answers to our questionnaire.

Threats: To evaluate the impact of the spiky ball, we measure the time from pressing the start button until the first successful grasp using the trials of the second block. The data is first filtered to remove outliers in overall completion time ($\pm 2\sigma$). Ninety-eight percent of the data is retained after filtering. A two-way repeated measures ANOVA was used with Greenhouse-Geisser sphericity corrections. The two independent variables used in the ANOVA were visualization and ball type (smooth or spiky). We found a significant effect of ball type ($F(1, 22)$, $p < 0.05$), with the spiky ball resulting in overall longer times (Figure 6 (a)), and a significant effect of visualization ($F(7, 154)$, $p < 0.05$), but no interaction effect between visualization and ball type. To check for significant differences between individual visualizations, the data was modeled using a linear mixed effects model with completion time as the dependent variable, and visualization and ball type as predictors. A least-square means pairwise comparison with Tukey adjustment showed that the time to grasp was significantly higher with the spiky ball ($p < 0.05$) for OHTR.



(a) (b) (c)

Figure 6: Ownership results: (a) Time taken to grasp the ball with and without threat (spiky ball). (b) Completion times for block 1 (no saw) and block 3 (saw). (c) Mean response to ownership questions O1, O4, O6 and the combined ownership questions.

Table 1: Questionnaire items. Statements with significant main effects of visualization are highlighted in gray.

Concept	Statement
Ownership	O1: I felt as if the virtual hands were part of my body.
	O2: It sometimes seemed like my own hands came into contact with the virtual object.
	O3: I thought that the virtual hands could be harmed by a virtual danger.
	O4: I felt that my real body was endangered during the experiment.
	O5: I felt that my real hand was endangered during the experiment.
	O6: I anticipated feeling pain from the spinning saw on the screen.
	O7: I tried to avoid the virtual saw while performing the task.
Agency	A1: I felt as if I can control movements of the virtual hands.
	A2: I felt as if the virtual hands moved just like I wanted them to, as if they were obeying my will.

To analyze the effect of the saw in the third block, we compare the task completion time in blocks 1 (no threat) and 3 (saw). One measure was removed because our software crashed during the trial. A linear mixed effects model showed a main effect of the saw on completion time (Figure 6 (b)). A post-hoc (least-square means pairwise comparison with Tukey adjustment) showed that for IH, OH, OHRA, and DH the time changed significantly.

Questionnaire: Seven ownership and two agency questions were presented after completion of the third block for each grasping feedback condition (see Table 1). The Friedman rank test was used to test for effects of grasping feedback on responses per question. If significant effects were found, a post-hoc pairwise Wilcoxon test was used and differences with a p-value of $p < 0.01$ were reported. A significant effect of visualization was found ($p < 0.05$) for questions O1, O4, and O6 in Table 1, see Figure 6 (c). The post-hoc test for O1 showed significant differences between DH/2H, DH/IHTR, and DH/OH. For O4, DH was rated significantly lower than OH, and for O6, OH was rated significantly lower than OHRA.

The Friedman test showed a significant effect ($p < 0.005$) of visualization for the averaged responses to the seven ownership questions. The post-hoc Wilcoxon test showed significantly different responses between DH and OH.

There were no significant differences in responses to the two agency questions, suggesting that users felt an equally strong sense of control over the virtual hands for each condition.

Summary: These results clearly **do not support our hypotheses H3 and H4** that visualizing the tracked hand or adding the reactive affordance visualization results in a stronger sense of embodiment. On the contrary, based on the graphed results, any tendency points in the other direction with the OH and OHTR conditions indicating a larger sense of ownership. Participants did react to the threats, with IH, OH, OHTR, OHRA, and DH displaying significant increases for the time to grasp or completion time with the spiky ball or saw. Interestingly, DH was rated lowest in some of the ownership questions, which should not be surprising as the hand simply disappeared during the manipulation. However, the users still strongly reacted to the saw as a threat.

4.3 Preference

Participants rank all conditions in order of preference once at the beginning, pre-experiment, so that their ranking is only based on the appearance of the visualization and not on its use, and once at the end, post-experiment after having experienced each condition in practice. During the second ranking, participants were able to interact with the ball using the selected visualization.

In general, users preferred the visualizations in the outer hand group (OH, OHTR, OHRA) over those in the inner hand group (IH, IHTR, IHRA), as hypothesized. 2H was ranked between OH and IH on average, and DH was consistently among the least preferred in both preference measures, with some participants reporting that it was "jarring", though a few ranked it more favorably after using it.

Although the differences in rankings for the visualizations were more pronounced in the pre-experiment ranking (see Figure 7 (right)), a Wilcoxon rank test showed that they did not change significantly after the participants used each visualization, suggesting a preference for natural looking interactions despite the

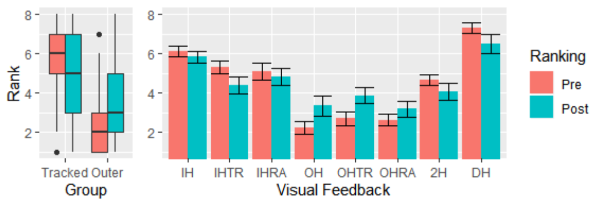


Figure 7: Pre and post experiment rankings for each visualization group (left). Average pre and post experiment rankings for each visualization (right). 1 is the most preferred.

performance penalty. Specifically looking at the inner and outer hand conditions (Figure 7 (left)), a Friedman test showed that there were significant differences in their averaged rank in both the pre-experiment ($p < 0.0001$) and post-experiment ranking ($p < 0.005$). However, the difference is less pronounced in the post-experiment condition.

The resulting rankings of the visualizations **confirm our hypothesis H5** that visualizations without intersections will be visually preferred. However, after experiencing each condition in practice, this difference was not as pronounced anymore.

5 DISCUSSION AND LIMITATIONS

In our experiment, we were able to confirm our assumptions H1 and H5, but we could not find evidence for H2, H3, and H4.

As hypothesized, we found that the visualizations in the tracked hand group resulted in better performance than those in the outer hand group. This result might be due to the clear visual feedback on their hand posture and the finger-object intersections. The grasping performance for 2H fell between the tracked hand and outer hand visualizations, which might indicate that the outer hand representation could be distracting. This result partially agrees with Prachyabrued and Borst's [2014] findings for performance: IH performed best, OH was worst, and 2H was between the two. Regarding fingertip depth, our results suggest that the outer hand visualizations caused a looser grip than the tracked hand visualizations, whereas Prachyabrued and Borst found the opposite result. One reason for that discrepancy could be that in their experiment, participants wait to release the object. During that wait time they might adjust their hand posture based on the visual feedback and tighten their grip for OH as it seems that the fingers are still on the surface. Interestingly, in our study, adding transparency to indicate the exact timing of the grasp and release did not improve performance.

For ownership, the time to grasp the ball was longer with the spiky ball and the completion time was longer for the saw, indicating hesitation to pick up the ball or to move past the saw. This result is similar to the results in Argelaguet et al. [2016], where the barbed wire and flame resulted in longer task completion times. Hesitation could be due to an elevated sense of endangerment. However, other factors may include surprise (though we expect that this effect diminishes due to repeated exposure), conditioning from games to avoid dangerous virtual obstacles, or planning of a different grip. The questionnaire responses and results from the spiky ball threat point to OH and OHTR resulting in the highest levels of ownership. DH resulted in the lowest level of ownership, likely because the hand is not visible while the user is holding the ball. Still, users reacted to the saw in that condition. Users felt a strong sense of agency over the virtual hand with each of the visualizations, most likely due to the high fidelity hand tracking used.

User preference was found to align with our hypothesis that the visualizations in the outer hand group will be preferred, which is also in line with the results from Prachyabrued and Borst [2014]. Participants ranked the feedback conditions at the beginning and at the end of the experiment. The outer hand group was preferred in both cases, even though performance was better for the tracked hand group. This suggests that users prefer more realistic interactions even if this means a loss of performance.

Limitations to this experiment include the repetitive task, the within subjects design, and the short amount of time spent with each condition (typically 1-2 minutes), which might have led to weaker differences between conditions and increasing familiarity with the virtual threats. Other limitations include the relative ease of the task and the use of a sphere as the only object to grasp. Post-experiment preference may have differed more from pre-experiment preference if the task were more difficult. Though previous studies have indicated that ownership can be established over unrealistic virtual hands and even non-corporeal objects, the use of a robot avatar in this study could have reduced the sense of danger from the threats.

6 CONCLUSION

We tested the performance, sense of ownership, and user preference of eight visual feedback techniques for virtual grasping. Our main results are:

- Visualizing the tracked hand (IH group) results in better performance than if the outer hand (OH group) is displayed.
- We could not confirm that visualizing the tracked hand (IH) or adding reactive affordance (IHRA and OHRA) would increase the level of ownership.
- Visualizations that prevent hand-object interpenetrations (OH group) were preferred, despite their lower performance.
- Not showing the hand at all (DH) was the least advantageous condition in preference and ownership.

Based on these results, we recommend OH or OHTR in applications where user preference and ownership are prioritized. If performance is essential, a tracked hand visualizations should be used, with IHTR being the most preferred among them. 2H can be a good compromise when both performance and preference are important. In general, we would recommend against a condition where the hand disappears such as DH, even if it is used in practice. It was preferred less than the others and resulted in the lowest sense of ownership.

Future work might use further measures for ownership, such as the galvanic skin response, or test further hand appearances such as a very realistic hand as this might affect how the visualizations are perceived.

ACKNOWLEDGMENTS

We would like to thank Cesare Parise for his suggestions concerning the reactive affordance conditions. This material is based in part upon work supported by the National Science Foundation under Grant Number IIS-1652210.

REFERENCES

Ferran Argelaguet and Carlos Andújar. 2013. A survey of 3D object selection techniques for virtual environments. *Computers & Graphics* 37, 3 (May 2013), 121–136. https://doi.org/10.1016/j.cag.2012.12.003

Ferran Argelaguet, Ludovic Hoyet, and Anatole Trico, Michaël & Lecuyer. 2016. The role of interaction in virtual embodiment: Effects of the virtual hand representation. In *2016 IEEE Virtual Reality*. 3–10. https://doi.org/10.1109/VR.2016.7504682

Christoph W. Borst and Arun P. Indugula. 2005. Realistic Virtual Grasping. In *IEEE Proceedings. VR 2005. Virtual Reality, 2005.* 91–98. https://doi.org/10.1109/VR.2005.1492758

Doug A. Bowman and Larry F. Hodges. 1997. An Evaluation of Techniques for Grabbing and Manipulating Remote Objects in Immersive Virtual Environments. In *Proceedings of the 1997 Symposium on Interactive 3D Graphics (I3D '97)*. ACM, 35–ff. https://doi.org/10.1145/253284.253301

Doug A. Bowman, Ryan P. McMahan, and Eric D. Ragan. 2012. Questioning Naturalism in 3D User Interfaces. *Commun. ACM* 55, 9 (Sept. 2012), 78–88. https://doi.org/10.1145/2330667.2330687

Andreas Geiger, Imke Bewersdorf, Elisabeth Brandenburg, and Rainer Stark. 2018. Visual Feedback for Grasping in Virtual Reality Environments for an Interface to Instruct Digital Human Models. In *Advances in Usability and User Experience*. Springer International Publishing, 228–239. https://doi.org/10.1007/978-3-319-60492-3_22

Shangchen Han, Beibei Liu, Robert Wang, Yuting Ye, Christopher D. Twigg, and Kenrick Kin. 2018. Online Optical Marker-based Hand Tracking with Deep Labels. *ACM Trans. Graph.* 37, 4, Article 166 (July 2018), 10 pages. https://doi.org/10.1145/3197517.3201399

Konstantina Kilteni, Raphaela Groten, and Mel Slater. 2012a. The Sense of Embodiment in Virtual Reality. *Presence: Teleoperators & Virtual Environments* 21, 4 (Dec. 2012), 373–387. https://doi.org/10.1162/PRES_a_00124

Konstantina Kilteni, Jean-Marie Normand, Maria V. Sanchez-Vives, and Mel Slater. 2012b. Extending Body Space in Immersive Virtual Reality: A Very Long Arm Illusion. *PloS ONE* 7 (July 2012), 1–15. https://doi.org/10.1371/journal.pone.0040876

Meng Chun Lam, Haslina Arshad, Anton Satria Prabuwono, Siok Yee Tan, and S. M. M. Kahaki. 2018. Interaction techniques in desktop virtual environment: the study of visual feedback and precise manipulation method. *Multimedia Tools and Applications* 77, 13 (Jul 2018), 16367–16398. https://doi.org/10.1007/s11042-017-5205-9

Lorraine Lin, Alexandra Adkins, Yu Sun, Andrew Robb, Yuting Ye, Massimiliano Di Luca, and Sophie Jörg. 2019. The Effect of Hand Size and Interaction Modality on the Virtual Hand Illusion. In *2019 IEEE Virtual Reality Conference*.

Lorraine Lin and Sophie Jörg. 2016. Need a Hand?: How Appearance Affects the Virtual Hand Illusion. In *Proceedings of the ACM Symposium on Applied Perception (SAP '16)*. ACM, 69–76. https://doi.org/10.1145/2931002.2931006

C. Karen Liu. 2008. Synthesis of Interactive Hand Manipulation. In *Proceedings of the 2008 ACM SIGGRAPH/Eurographics Symposium on Computer Animation (SCA '08)*. Eurographics Association, Aire-la-Ville, Switzerland, Switzerland, 163–171. http://dl.acm.org/citation.cfm?id=1632592.1632616

Matthew Longo, Friederike Schüür, Marjolein P M Kammers, Manos Tsakiris, and Patrick Haggard. 2008. What is embodiment? A psychometric approach. *Cognition* 107 (July 2008), 978–98. https://doi.org/10.1016/j.cognition.2007.12.004

Ke Ma and Bernhard Hommel. 2013. The virtual-hand illusion: effects of impact and threat on perceived ownership and affective resonance. *Frontiers in Psychology* 4 (Sept. 2013), 604. https://doi.org/10.3389/fpsyg.2013.00604

Ke Ma and Bernhard Hommel. 2015a. Body-ownership for actively operated non-corporeal objects. *Consciousness and Cognition* 36 (June 2015), 75 – 86. https://doi.org/10.1016/j.concog.2015.06.003

Ke Ma and Bernhard Hommel. 2015b. The role of agency for perceived ownership in the virtual hand illusion. *Consciousness and Cognition* 36 (July 2015), 277 – 288. https://doi.org/10.1016/j.concog.2015.07.008

Ryan P. McMahan, Doug A. Bowman, David J. Zielinski, and Rachael B. Brady. 2012. Evaluating Display Fidelity and Interaction Fidelity in a Virtual Reality Game. *IEEE Transactions on Visualization and Computer Graphics* 18, 4 (April 2012), 626–633. https://doi.org/10.1109/TVCG.2012.43

Owlchemy Labs. 2016. Job Simulator: The 2050 Archives. Oculus Rift version.

Nancy S. Pollard and Victor B. Zordan. 2005. Physically Based Grasping Control from Example. In *Proceedings of the 2005 ACM SIGGRAPH/Eurographics Symposium on Computer Animation (SCA '05)*. ACM, 311–318. https://doi.org/10.1145/1073368.1073413

Ivan Poupyrev, Suzanne Weghorst, Mark Billinghurst, and Tadao Ichikawa. 1997. A Framework and Testbed for Studying Manipulation Techniques for Immersive VR. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology (VRST '97)*. ACM, 21–28. https://doi.org/10.1145/261135.261141

Mores Prachyabrued and Christoph W. Borst. 2012. Visual interpenetration tradeoffs in whole-hand virtual grasping. In *2012 IEEE Symposium on 3D User Interfaces (3DUI)*. 39–42. https://doi.org/10.1109/3DUI.2012.6184182

Mores Prachyabrued and Christoph W. Borst. 2013. Effects and Optimization of Visual-Proprioceptive Discrepancy Reduction for Virtual Grasping. *2013 IEEE Symposium on 3D User Interfaces (3DUI)*, 11–14. https://doi.org/10.1109/3DUI.2013.6550190

Mores Prachyabrued and Christoph W. Borst. 2014. Visual Feedback for Virtual Grasping. *2014 IEEE Symposium on 3D User Interfaces (3DUI)* (March 2014), 19–26. https://doi.org/10.1109/3DUI.2014.6798835

Schell Games. 2016. I Expect You To Die. Oculus Rift version.

Martin Schubert and Barrett Fox. 2017. Interaction Sprints at Leap Motion: Exploring the Hand-Object Boundary. http://blog.leapmotion.com/interaction-sprint-exploring-the-hand-object-boundary/. Accessed July 15, 2019.

Mel Slater, Daniel Pérez-Marcos, H. Henrik Ehrsson, and Maria V. Sanchez-Vives. 2008. Towards a Digital Body: The Virtual Arm Illusion. *Frontiers in Human Neuroscience* 2 (Aug. 2008), 1499 – 1506. https://doi.org/10.3389/neuro.09.006.2008

Spyros Vosinakis and Panayiotis Koutsabasis. 2018. Evaluation of visual feedback techniques for virtual grasping with bare hands using Leap Motion and Oculus Rift. *Virtual Reality* 22, 1 (01 Mar 2018), 47–62. https://doi.org/10.1007/s10055-017-0313-4

Ye Yuan and Anthony Steed. 2010. Is the Rubber Hand Illusion Induced by Immersive Virtual Reality?. In *Proceedings of the 2010 IEEE Virtual Reality Conference (VR '10)*. IEEE Computer Society, Washington, DC, USA, 95–102. https://doi.org/10.1109/VR.2010.5444807

Wenping Zhao, Jianjie Zhang, Jianyuan Min, and Jinxiang Chai. 2013. Robust Realtime Physics-based Motion Control for Human Grasping. *ACM Trans. Graph.* 32, 6, Article 207 (Nov. 2013), 12 pages. https://doi.org/10.1145/2508363.2508412