Temporal Availability of Ebbinghaus Illusions on Perceiving and Interacting with 3D Objects in a Contextual Virtual Environment

Russell Todd*
3D Interactive Realities Lab
University of Wyoming

Qin Zhu†
UW PACE (Perception-Action-Cerebral-Executive) Lab
University of Wyoming

Amy Banić‡
3D Interactive Realities Lab
University of Wyoming

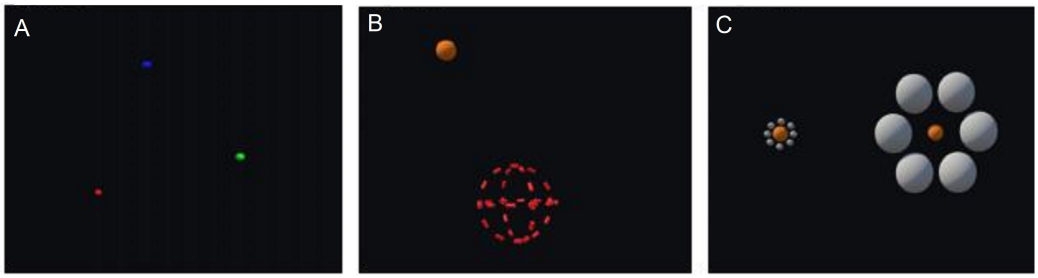

Figure 1: Experimental Environment: A) Three virtual Spheres represent 3D positions of hand grasp configuration- forefinger (blue), thumb (green), and wrist (red) sensors; B) Home Wireframe Sphere and Target Sphere; C) Ebbinghaus Illusion induced by smaller 3D contextual spheres (left) and larger 3D contextual spheres (right)

Abstract

Contextual illusions, such as the Ebbinghaus Illusion, can be potentially used to improve or hinder reach-to-grasp interaction in a virtual environment by affecting the perception of object size and the action. However, the illusion effect has only been evaluated using 2D objects like discs or annuluses, and limited research has been conducted in a virtual environment. Moreover, it remains unknown how the sudden, or dynamic, change of surrounding features will impact the perception and then the action towards the object. In this paper, we conducted a series of experiments to evaluate the effects of 3D Ebbinghaus illusion with dynamic surrounding features on the task of reaching to grasp a 3D object in an immersive virtual environment. An innovative 3D perceptual judgment task revealed that the static 3D illusion affected the perceived size of the 3D object. Then, we experimentally manipulated the visual gain and loss of the 3D contextual inducers, the participant's virtual hand, and the entire 3D contextual object. Results revealed that the depth error (error in depth of reaching action) was influenced by a dynamic change in the size of the inducers, the fine adjustment of grasp was dependent on the visual presence of the virtual hand and vision of 3D contextual object was required for the reaching and grasping movements. These results will benefit the understanding of reach-to-grasp interactions in immersive virtual environments and can improve interaction design.

Keywords: 3d perceptual judgement task, reach-to-grasp, Ebbinghaus, illusion, perception, 3D Contextual Objects, temporal availability, dynamic inducers, static inducers, tracking, virtual reality

Index Terms: Human-centered computing [Interaction design]: Interaction design theory— concepts and paradigms Human-centered computing [Human computer interaction (HCI)]: Interaction paradigms—Virtual reality

*e-mail: rtodd3@uwyo.edu
†e-mail: qzhu1@uwyo.edu
‡e-mail: abanic@cs.uwyo.edu

1 Introduction

Reaching to grasp an object situated in a context is commonly involved in daily living and occupational tasks (e.g. reaching to grasp an apple from a fruit basket, a coffee mug from a shelf, or a chair in a classroom). It is also an important and common task for many immersive virtual reality applications. Previous studies have shown that the reach-to-grasp movement is governed by both planning and control mechanisms, and the actor is adept at adjusting the reaching and grasping movements based on the perception of the to-be-grasped object [11] [2]. A user in an immersive virtual environment (VE) must adjust the reaching and grasping movements based on the perception of the to-be-grasped 3D object as well. The perception of the object is contextual as demonstrated by the Ebbinghaus illusion in which the perceived size of the center circle varies depending on the size of surrounding inducers (as can be seen in Figure 1, C). It has been shown that the Ebbinghaus illusion affects visual perception of the object size as well as the actions to interact with the object [9] [13] However, the illusion effect has only been evaluated using 2D objects like discs or annuluses, and it remained unknown if the illusion would persist with 3D objects and impact the interaction with 3D objects as well since size perception and depth perception are both required. In addition, the illusion effect has only been studied using constant configurations of surrounding features, and it remains unknown how the sudden change of surrounding features (the context) would impact the adjustment of the movement during interaction with the object and could impact 3D UI interaction performance. Thinking of an implementation is often the next step after learning about a new phenomenon, but most objects people use are three dimensional which lay outside the scope of these previous studies. Would this effect extend to 3D objects or could the effect differ due to additional size/depth cues like, to name just one example, the shading on a 3D object resulting from light on its surface? Using integrated motion tracking and immersive virtual reality (VR) technologies, researchers can monitor and assess the participant's perception and action upon 3D objects in real time. First, we devised a novel judgement task to determine if Ebbinghaus illusion persists in an immersive VR environment. In this paper we present the results of three experiments where we manipulated the contextual features of the 3D object to determine the static and dynamic effects of the 3D Ebbinghaus illusion on the reach-to-grasp

2021 IEEE Virtual Reality and 3D User Interfaces (VR) | 978-1-6654-1838-6/21/$31.00 ©2021 IEEE | DOI: 10.1109/VR50410.2021.00109

2642-5254/21/$31.00 ©2021 IEEE
DOI 10.1109/VR50410.2021.00109

movement in an immersive virtual environment. Finally, we manipulated the vision of the virtual hand and the entire contextual object to determine their respective roles in reach-to-grasp performance.

2 Related Works

The Ebbinghaus illusion has been known as evidence to show that a human's visual perception of object size is contextually influenced by the surrounding features of the object. Typically, an object would be perceived as larger when it is surrounded by smaller objects and smaller when surrounded by larger objects, although factors like age, environment and culture may affect the magnitude of the illusion effect [7] [5]. Would such a visual illusion affect the visually guided actions used to interact with the object? The answer to this question remains undetermined due to contradictory evidence from previous studies. In early studies examining the illusion effect on prehension and grasping behavior used to interact with the 2-D pictorial elements [1] [12] [17], a small or no illusion effect on action was reported, which leads to the dissociation theory of perception and action with respect to interaction with 2-D contextual objects [13] and two separate neural pathways responsible for processing visual information [16]. However, emerging studies challenged the dissociation theory by showing that the interactive action with the contextual object was affected by the illusion as much as the perception of the object would. Van Donkelaar [18] asked his participants to point rapidly toward a pictorial target circle surrounded by smaller or larger circles and found out that movement time was longer and less accurate when pointing toward a perceptually smaller target than toward a perceptually larger target. Franz et al. [9] examined grasping performance when the to-be-grasped annulus was presented in the pictorial context to induce Ebbinghaus illusion, and reported that the maximum aperture shaped between the fingers for grasping was reflective of the perceived size of the object in the illusion. In a study requiring participants to tap rapidly between two pictorial target circles surrounded by smaller or larger circles [2], a faster and more accurate tap was seen for the perceptually larger target.

Recently, the 2-D Ebbinghaus illusion effect has been investigated in an immersive 3D virtual environment [8]. Participants wore a HTC Vive Head-Mounted Display and saw contextual 2-D discs floating in a virtual room environment. They were asked to judge the size and distance of the center disc using a mouse and a controller. The results showed that size judgements were symmetrically affected by the illusion, where the center disc was perceived as larger with smaller inducers around and smaller with larger inducers around. The distance judgments were asymmetrically affected by the illusion, where the center disc was judged to be closer with surrounding smaller inducers but not different from surrounding larger inducers.

Without contextual features to induce the illusion, the performance of reaching to grasp a 3D object was recently examined and compared between physical and haptic-free virtual environments [10]. Furmanek et al. studied the reach-to-grasp kinematics of healthy subjects in a haptic-free virtual environment. The subjects were instructed to reach-to-grasp-to-lift objects of 3 different sizes set at 3 separate distances from the starting position. This task was performed both in a physical environment and a virtual environment so that the motion performances could be compared. As there was no haptic feedback in the virtual environment, they made use of a custom collision detection algorithm to establish participant contact with the virtual environment. After analysis, Furmanek et al. found that the kinematic profiles of both the transport velocity and grasp aperture were strongly correlated, with the most prominent difference being a prolonged closure phase of the reach to grasp movement in the virtual environment. The results showed that the coordination of reaching velocity and grasp aperture was preserved in a virtual environment as much as in a physical environment, suggesting that haptic-free virtual environments may be a useful platform to study reach-to-grasp movements, in which the visual properties of the 3D object can be systematically manipulated.

In sum, the effects of the Ebbinghaus illusion on perception and action have been examined using 2-D pictorial objects in both physical and virtual environments, however, it remained unknown whether the illusion effects would continue to affect people' perception and interaction with 3D objects, which is important knowledge given that we are living in a 3D world. In addition, in evaluating the illusion effect on interactive actions, no one has examined how temporal availability of visual information such as the contextual features of the 3D object and vision of the hand could impact the planning and control of the action. Such knowledge will help determine what and when visual information is used for effective interaction.

3 Experiments

3.1 Apparatus

For all three studies, the immersive virtual environment (IVE) was developed in the Unity Engine and displayed using an HTC Vive virtual reality head-mounted-display (HMD). The Vive was chosen as it boasted the best commercially available VR display at the time so that we could avoid stereo display deficiencies as much as possible within budget [3] [14]. The internal tracking of the Vive system was used for head tracking and tracking the non-dominant hand using a HTC Vive controller. The participant's dominate hand was tracked using a trakSTAR electromagnetic (EM) tracking device. The EM sensors were mounted upon a spandex glove via Velcro straps to be on top of the fingernails of the forefinger and thumb of the dominant hand when the glove was worn by the participant. A third EM sensor was mounted on a Velcro strap attached to the cuff of the glove so that the sensor would sit on the midpoint between the radius and ulna bone heads of the participant's wrist when the strap was fastened. (Figure 2, B & C).The EM tracking data was simultaneously streamed by a MATLAB program to Unity, recorded by a Unity C# program, and exported as a csv file upon termination.

3.2 Participants

Prior to beginning, all participants were informed of the procedures involved in the experiment. The procedures were reviewed by the University of Wyoming Institutional Review Board. If participants then agreed to participate, they were screened for normal or corrected-to-normal vision by verbal report. Participants were all college students between the age of 21 and 27 attending the University of Wyoming. Participants were offered extra credit as an incentive for their participation in this experiment.

Experiment 1: Nineteen participants (9 male, 2 participants omitted due to incomplete data) were recruited from the population of Laramie, WY at the University of Wyoming.

Experiment 2: Seventeen participants (9 male) were recruited from the population of Laramie, WY at the University of Wyoming.

Experiment 3: Thirteen participants (5 male) were recruited from the population of Laramie, WY at the University of Wyoming.

3.3 Virtual Environment

The virtual environment was a featureless space made to be as close to black as possible inhabited only by the user and the trial structures. Only the user's hand representation and trial structures had lighting enabled. This was chosen so that the participant would have as few references as possible when judging the size of the target. The trial structures consisted of the target, a 'home' sphere, and the inducers (aka surrounding features or Ebbinghaus structures) necessary for the Ebbinghaus Illusion. The target was an orange sphere 3.56cm in diameter. The small inducer spheres were 1.65cm in diameter and the large inducer spheres were 8.95cm in diameter. The home sphere was a sphere constructed of a wire-frame of dotted lines. The vertical distance between the home and target spheres was 38.875cm. This structure initially appeared red if the participant's fingers were outside it and then turned to green if the participant's fingers entered and remained inside it. Upon the departure of fingers from the home sphere, the home sphere would disappear. The Ebbinghaus inducers consisted of 6 large (8.95cm diameter) white spheres and 8 small (1.65cm diameter) white spheres as shown in the reference image. These inducers were never active at the same time. The large inducers would make the target appear smaller while the small inducers would make the target appear larger.

Experiment 1	Conditions				
Block 1: Static Inducers	Control	Small	Large		
Block 2: Dynamic Inducers (Appear/Disappear)	Control	Small Appear	Small Disappear	Large Appear	Large Disappear
Block 3: Dynamic Inducers (Swapping)	Control	Small to Large	Large to Small		
Block 4: Sphere Creation	Control	Small	Large		
Experiment 2					
Occluded Early	Exp. 1 Block 1	Exp. 1 Block 2	Exp. 1 Block 3		
Occluded Late	Exp. 1 Block 1	Exp. 1 Block 2	Exp. 1 Block 3		
Occluded Entirely	Exp. 1 Block 1	Exp. 1 Block 2	Exp. 1 Block 3		
Experiment 3					
Perturbed Early	Exp. 1 Block 1	Exp. 1 Block 2	Exp. 1 Block 3		
Perturbed Late	Exp. 1 Block 1	Exp. 1 Block 2	Exp. 1 Block 3		

Figure 2: Complete Experimental Conditions Table

4 Procedure for All Experiments

4.1 Experiment 1:

4.1.1 Experimental Design:

The study was intended to determine if the Ebbinghaus illusion effect persisted with 3D objects and how the action of reaching to grasp the 3D object would be affected by the static and dynamic illusion in the immersive VR environment. In the Judgment blocks, participants were asked to create a 3D object on the right side of visual field using Thumb and Forefinger to match the size of a 3D object displayed in equal depth on the left side of their visual field with or without contextual inducers. Three conditions were tested: the Control condition in which the 3D object appeared with no surrounding inducers; the Small Inducer condition in which the 3D object appeared with small inducers in the surrounding; the Large Inducer condition in which the 3D object appeared with large inducers in the surrounding. Each condition was repeated for 10 times but in a random order. In the Action blocks, participants were asked to reach to grasp the 3D object displayed in a fixed distance in front as natural as possible. Both the 3D object and virtual hand were kept visible to the participants. Block 1 consisted of 3 conditions in which the contextual inducers, if present, remained still with the target: Control (only target), Small Inducer (target surrounded by small inducers), and Large Inducer (target surrounded by large inducers) with each condition repeated for 10 times in a random order, a total of 30 trials. Block 2 consisted of 5 conditions in which the contextual inducers either appeared or disappeared around the target in the midway of reaching toward the target: Control, Small Inducer Appear, Large Inducer Appear, Small Inducer Disappear, Large Inducer Disappear, with each condition repeated for 10 times in a random order, a total of 50 trials. Block 3 consisted of 3 conditions in which the contextual inducers swapped in size in the midway of reaching toward the target: Control, Large-to-Small, Small-to-Large, with each condition repeated for 10 times in a random order, a total of 30 trials. The Large-to-Small refers to the condition where the 3D object was initially surrounded by Large inducers, but the inducers were switched to be Small in the midway of reaching toward the target, which was opposite for Small-to-Large condition.

4.1.2 Procedure of Perceptual Judgment Block:

The participant was given a moment to orient themselves. The MATLAB program controlling the collection and communication of sensor position data to Unity was started, which signaled the simulation to start recording. The forefinger sensor and HTC Vive controller were used to calibrate the positions of the EM tracker emitter and the HTC Vive virtual space. The position for trial structures was also initialized to a position near the eye level of the participant. Participant placed fingertips within the home sphere which served as a start position. Participant completed task, which was for them to reach out with their tracked hand into the 'creation space' and create a sphere identical in size to the target sphere to the front and left of them, which may or may not be surrounded by Ebbinghaus features. When their hand was in the correct position, an orange sphere came into being between the tips of their forefinger and thumb, with the distance between them being the diameter of the sphere. They could adjust the size of the sphere simply by widening or narrowing the gap between their finger and thumb. When they felt that the created sphere was the correct size, they pressed a button on the Vive controller in their offhand which set the created ball at its current size. With that completed, they returned to the home sphere. At that time, the facilitator clicked a button on their own controller to begin the next trial. During experiments, operator sat behind and to the side of the participant to keep the operator's controller, which would be visible in the IVE, out of view of the participant. When the experiment was complete (30 trials) the simulation ended. The participant then began the next experimental block, repeating this exercise until all experiments were complete.

4.1.3 Procedure of Action Blocks (Blocks 1-3):

Participant arrived for testing and was informed that the main task of experiment was to reach-to-grasp an orange spherical target that would appear in their visual field. Participant entered testing area and had EM sensors attached to forefinger, thumb, and wrist by wearing the spandex glove on his/her dominant hand before putting on the HTC Vive head-mounted device. A training VR environment was provided for the participant to become acquainted with the system and its basic functionality. During training, a target appeared at one of 9 positions in front of the participant simulating the 9 directions that the participant could reach to grasp the target. To be ready for the trial, the participant was told to close the hand with attached forefinger and thumb, and then keep the finger markers within the home sphere so that the home sphere turned to green. Upon seeing the target, the participant would reach to grasp the target by moving the finger markers toward and closer to the target until the two finger markers are perceived to touch the edge of the target spanning in a length of the diameter of the target. To confirm that the target was fully grasped, the participant would click the trigger of the controller on the other hand. Once clicked, the current target disappeared and the next target would appear, so on and forth. The home sphere would disappear once the finger markers left home, and it would become visible again after the click. During training, participants may receive instructions how to better perform the requested task. Example instructions include to do the motion without slowing down too much so that the reach-to-grasp movement would be based on the immediate perception of the target. The training continues until the participant is prepared to begin the experiments. At this point the Unity simulation is started with the appropriate experiment selected.

The participant is given a moment to orient themselves. The MATLAB program controlling the collection and communication of

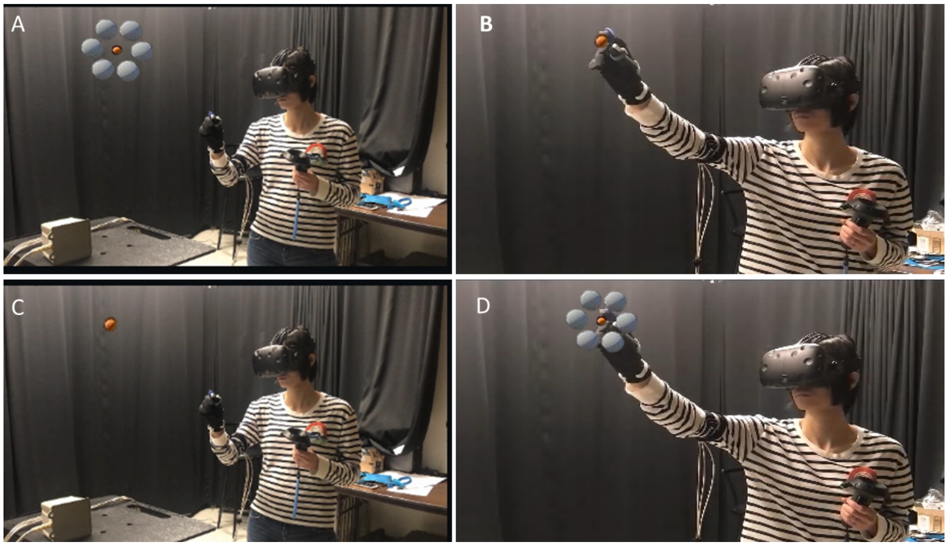

Figure 3: Example of experimental conditions where temporal availability of the surrounding contextual spheres changes during the task. From A to B: Inducers visible when participant begins task (A), but disappear as hand is closer to 3D object. From C to D: Inducers not visible when participant begins task (A), but appear as hand is closer to 3D object.

sensor position data to Unity is started, which signals the simulation to start recording. The forefinger sensor and HTC Vive controller are used to calibrate the positions of the EM tracker emitter and the HTC Vive virtual space. The forefinger sensor is held in the center of the ring of the Vive controller. In this way we know the relative positions of the controller and sensor to each other. The controller and finger are moved to three different positions and a position measurement is taken. Using the three sets of points, the coordinate systems of the VR tracking system and the EM tracking system are aligned. The position for trial structures is also initialized to a position near the eye level of the participant. Participant places fingertips within the home sphere which serves as a start position. Participant completes task, which is simply for them to reach out with their tracked hand as if they were going to grab the target ball. When they feel that their hand is in the correct position, they press a button on the Vive controller in their offhand which 'grasps' the target ball. They then can drag it off to the side and release it. With that completed, they return to the home sphere. At that time, the operator administering the experiment will click a button on their own controller to begin the next trial. When the experiment is completed (30 or 50 trials) the simulation ends. The participant then begins the next experimental block, repeating this exercise until all experiments are complete.

4.2 Experiment 2:

4.2.1 Experimental Design:

As a follow-up of Experiment 1, Experiment 2 was intended to determine if the visual gain or loss of virtual hand would affect the reach-to-grasp performance with the presence of static and dynamic 3D Ebbinghaus illusion. Thus, the vision of the virtual hand was occluded at early, late or entire phase of the reach-to-grasp motion. Consequently, the action blocks of Experiment 1 were tripled.

4.2.2 Experiment 2 Procedure for Blocks 1-3:

The procedure is the same as the one used in the Experiment 1 action blocks, with the following changes. The virtual 'hand' in all experiments consists of the three multicolored dots which markers the sensors for the forefinger, thumb, and wrist. When 'occluded' the three dots are made invisible in the virtual environment leaving the participant with only their proprioception to inform them of the position of their hand. The early phase, late phase, and entire phase are defined by two positions of interest in the virtual environment. The early phase begins when the participant's hand leaves their starting position, and the late phase begins when the participant's hand reaches the halfway point to the target. In an early condition, the hand is disappeared after they leave the home sphere and reappears when they reach the halfway point. In a late condition, the hand is visible until the hand reaches the halfway mark at which point the hand disappears. In the entire condition, the handle is invisible for both phases, visible only while inside the starting position.

4.3 Experiment 3:

4.3.1 Experimental Design:

Experiment 3 was intended to evaluate if the visual gain or loss of contextual 3D object would affect the reach-to-grasp performance with the presence of dynamic 3D Ebbinghaus illusion. Thus, the vision of the contextual 3D object was occluded at early or late phase of the reach-to-grasp motion, while the virtual hand remained visible. Consequently, the action blocks of Experiment 1 were doubled.

4.3.2 Experiment 3 Procedure for Blocks 1-3:

The procedure is the same as the one used in the Experiment 1 action blocks, with the following changes. The same timing of Early & Late in Experiment 2 is used, but instead of the hand disappearing, it is the 3D target with or without inducers that was occluded. As some of those objects are active in certain conditions (Ebbinghaus illusion inducers in the swapping condition, for example), a change was made to allow for an active change to occur. As block 1 includes no active changes, its behavior remains the same. In blocks 2 & 3 (Appear/Disappear & Swap), when the participant leaves the home position, the 3D objects setup as usual. If it is the early condition, then the initial inducers (large or small) will take its action when the participant leaves the home position (as usual) and then disappear after a 0.1 second delay. Explicitly, this would mean appearing and then disappearing after 0.1 seconds (block 2) or swapping from one to the other and then disappearing after 0.1 seconds (block 3). In the late condition, this change is omitted as the participant would already have time to view the initial conditions during the first half of the reaching motion.

4.4 Analysis

Final data analysis of experiments 1, 2, and 3 was completed using a custom built program in the Unity engine. Trial start and end was defined as the instant the virtual hand left the home sphere and the virtual hand grasped the target, respectively. In the judgment block of experiment 1, only the 3D coordinates of the sensors attached to the thumb and forefinger at the moment of clicking the button of the controller were recorded, and the distance between the two digits was calculated to compare with the true diameter of the target sphere to determine the perceived size of the target sphere. Events included were entering the home sphere, leaving the home sphere, reaching the halfway point, and grasping the target. All kinematic data was linearly interpolated and filtered using a 6Hz 4th order Butterworth filter. Data captured from the sensors attached to the forefinger, thumb, and wrist was used to calculate the following kinematic measures. Movement time (MT) is the difference in time between the start and end of every reach-to-grasp motion. Peak grasp aperture (PGA) is the maximum distance reached between the forefinger and the thumb over the course of the reach-to-grasp motion. The peak reaching velocity (PRV) is the maximum velocity achieved by the wrist sensor over the course of the movement. The relative time to PGA and PRV was calculated as the ratio between the time to PGA/PRV and the MT. The terminal grasp aperture (TGA) was the distance between the forefinger and thumb sensors at the instant the target was grasped. Terminal grasp error (TGE) was quantified by calculating the depth error (distance between X coordinates) and plane error (perpendicular plane to participant and distance between YZ coordinates) using the center position of the target and midpoint between the forefinger and thumb at the instant of grasping.

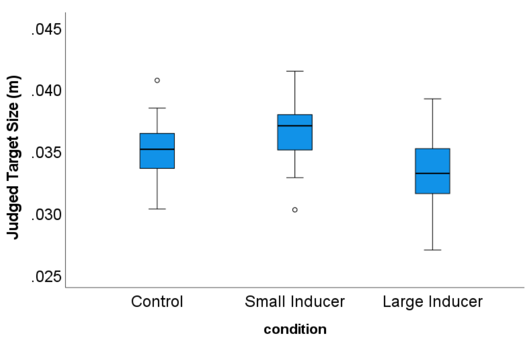

Figure 4: Experiment 1: Mean Judged Target Size Across Conditions: Control, Small Inducers, and Large Inducers

4.5 Statistics

First, all aforementioned variables (judgment and action) were calculated in each trial for each participant. Then, outliers were identified by boxplotting the 10 trials of data before being removed. Finally, the mean value of each variable was calculated by averaging across valid trials. Subsequently, Shapiro-Wilk test was performed on all these mean variables to examine the normality of data. For variables that are normally distributed, the repeated measures of analysis (rmANOVA) was performed, otherwise, the nonparametric Friedman test was performed instead to examine the effects of contextual conditions on perceiving and interacting with the 3D object. Specifically, in experiment 1, one-way (3 or 5 contextual conditions) rmANOVA or Friedman test was used to examine the effects of static illusion on perception and action, as well as the effects of dynamic illusion (appearing/disappearing or swapping inducers) on action. In experiment 2 and 3, one-way rmANOVA or Friedman test was used to examine additionally the effect of vision of hand or contextual object on action. In case of detecting significant main effect in rmANOVA, the post-hoc Tukey HSD test was used to determine the pair-wised difference. A multiple pairwise comparison using Nemenyi's procedure was used when significant effect was detected by Friedman test. The statistical significance level was set for all at $p < 0.05$. Statistical analysis was performed using SPSS (IBM Ver 22) running on a PC machine.

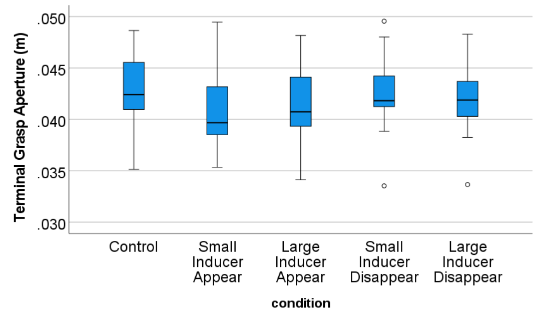

Figure 5: Experiment 1: Terminal Grasp Aperture Across Dynamic Conditions: Control, Small, Large X Appear/Disappear

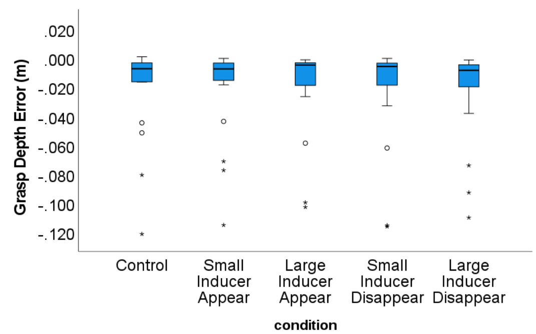

Figure 6: Experiment 1: Grasp Depth Error Across Dynamic Conditions: Control, Small, Large X Appear/Disappear

5 Results

5.1 Experiment 1:

5.1.1 Perceptual Judgment

The Ebbinghaus Illusion continued to affect the perception of object size in the 3D immersive VR environment, where the 3D object surrounded by small 3D inducers was perceived to be larger than control, and the one surrounded by large 3D inducers was perceived to be smaller than control. A Shapiro-Wilk test on this data was not significant, assuming a normal distribution. The one-way ANOVA on the mean judged target size showed a significant effect for condition ($F_{2,\ 36} = 20.03$, $p < 0.0001$). The post-hoc Tukey HSD test revealed that the judged target size was larger than control (M=0.035m, SD=0.002) ($p < 0.05$) when surrounded by small inducers (M=0.037m, SD=0.003), while it was smaller than control ($p < 0.05$) when surrounded by large inducers (M=0.033m, SD=0.003).

5.1.2 Static Illusion Effects on Reach-to-Grasp Movement

The illusion effect disappeared when people used their hands to interact with the object. The normality test showed that only TGAs were normally distributed, therefore, TGA was examined by rmANOVA, while all other kinematic measures were examined by Friedman test. The effect of contextual conditions was significant: $F_{2,36} = 9.14$, $p = 0.001$. The following post-hoc Tukey honest significant difference (HSD) test showed that the grasp aperture was smaller than control (M=0.06m, SD=0.02) ($p < 0.05$) when the object was surrounded by either large (M=0.06m, SD=0.01) or small (M=0.06m, SD=0.02) inducers with no difference detected between the latter two conditions ($p > 0.05$). Thus, the surrounding inducers constrained people in opening fingers to form an aperture to grasp the object. The Friedman test on other kinematic variables showed a significant difference among conditions only for movement time, $\chi^2(2) = 9.789$, p = 0.007. A multiple pairwise comparison using Nemenyi's procedure was followed and the results showed that it took longer to grasp the object when the object was surrounded by inducers than not surrounded.

5.1.3 Dynamic Illusion Effects on Reach-to-Grasp Movement

When inducers appeared or disappeared in the midway of reaching to grasp the object, the illusion effect remained absent and the presence of inducers only reduced the grasp aperture. Among all kinematic measures, only terminal grasp aperture (TGA) was normally distributed and the rmANOVA showed a significant effect for condition: $F_{4,72} = 5.83$, $p < 0.001$ for TGA. The following post-hoc Tukey HSD test showed that the grasp aperture was smaller than control ($p < 0.05$) whenever the object was surrounded by inducers earlier or later ($p > 0.05$). The Friedman test on other variables showed significant difference among conditions only for PGA, $\chi^2(4) = 13.558$, p = 0.009. As revealed by the Nemenyi's pairwise comparisons, PGA was significantly smaller when the inducers appeared earlier (in both disappear conditions). When inducers swapped in size in the midway of reaching to grasp the object, the illusion effect remained to be absent. The normality test showed that only TGA was normally distributed, and the rmANOVA revealed a significant effect for contextual conditions:($F_{2,36} = 4.97$, p = 0.012) (Figure 5). The following post-hoc Tukey HSD tests showed that the TGA was smaller (M=0.04m, SD=0.004) than control (M=0.043m, SD=0.003)($p < 0.05$) whenever there were inducers surrounding the object. The Friedman test on other variables showed significant difference only for DepthError, $\chi^2(2) = 9.579$, p = 0.008 (Figure 6). Nemenyi's multiple pairwise comparisons revealed there was more undershooting when the inducers changed from large to small.

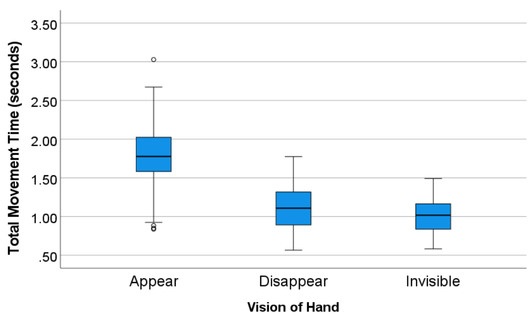

Figure 7: Experiment 2: Mean Total Movement Time Across Vision of Virtual Hand Conditions: Appear, Disappear, Invisible

5.2 Experiment 2:

Consistent with Experiment 1, the illusion effect remained absent and the presence of inducers only constrained the grasp aperture to make it smaller than control. However, the visual gain/loss of virtual hand significantly affected the kinematics of movement in all conditions of inducers. A Shapiro-Wilk test on all variables was significant, $p < 0.001$, therefore Friedman test was used to examine the effect of visual gain/loss of hand. Significant difference among hand conditions was found for the following variables: MT ($\chi^2(2) = 77.792$, $p < 0.001$), PGA ($\chi^2(2) = 28.125$, $p < 0.001$), and TGA ($\chi^2(2) = 62.042$, $p < 0.001$). The multiple pairwise comparison using

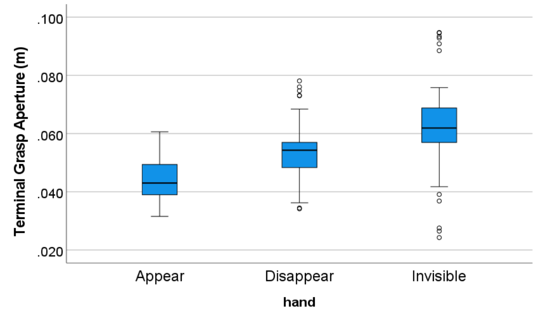

Figure 8: Experiment 2: Mean Terminal Grasp Aperture Across Vision of Virtual Hand Conditions: Appear, Disappear, Invisible

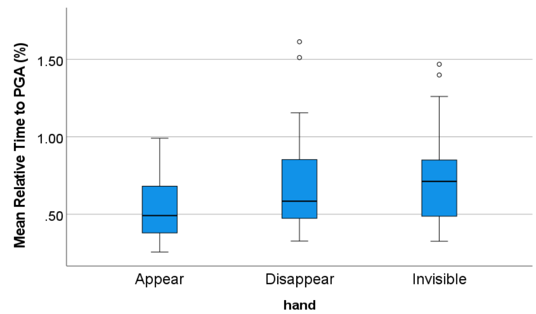

Figure 9: Experiment 2: Mean Relative Time to PGA Across Vision of Virtual Hand Conditions: Appear, Disappear, Invisible

Nemenyi's procedure revealed that: A) It took longer to complete the movement when hand was visible in the end of movement during grasping (M=2.19s, SD=0.62) than other conditions (Disappear, M=1.27s, SD=0.40) (Invisible, M=1.09s, SD=0.43); B) Compared to the condition where the hand was visible only in early or late phase of the movement (M-0.06m, SD=0.01), the peak aperture was greater when hand was invisible during the entire phase of movement (M=0.07m, SD=0.02). C) The terminal aperture was greatest when the hand was invisible during the entire movement phase, followed by the condition where hand was visible in early phase, and then the condition where the hand was visible in late phase. In sum, the visual gain of hand in the late phase not only increased movement time but also decreased grasp aperture, suggesting that vision of hand was required for final grasp adjustment for the object.

When inducers appeared or disappeared (Figures 7, 8, and 9), vision of virtual hand affected MT, PGA, TGA and Relative Time to PGA. The mANOVA was performed on the normally distributed PGA and TGA, and the results showed that the hand condition was significant for PGA ($F_{2,30} = 11.88$, $p < 0.0001$) and for TGA ($F_{2,30} = 20.57$, $p < 0.0001$). As revealed by the post-hoc Tukey HSD tests, the pattern of results for PGA and TGA repeated those seen earlier when inducers were constant. The Friedman test showed a significant difference among HAND conditions for both MT, $\chi^2(2) = 123.825$, $p < 0.001$, and Relative Time to PGA, $\chi^2(2) = 27.475$, $p < 0.001$. Nemenyi's procedure revealed that: 1) It took longer and longer as vision of hand was introduced from early to the end of reach-to-grasp movement; 2) the peak aperture occurred relatively earlier in the condition where the hand was visible in the end of movement (M=0.063m, SD=0.01) than the other two conditions where the hand was invisible in the end of movement (Disappear, M=0.066m, SD=0.02)(Invisable, M=0.076m, SD=0.01), further suggesting that vision of hand allowed participants to spend more time closing the fingers when grasping the object.

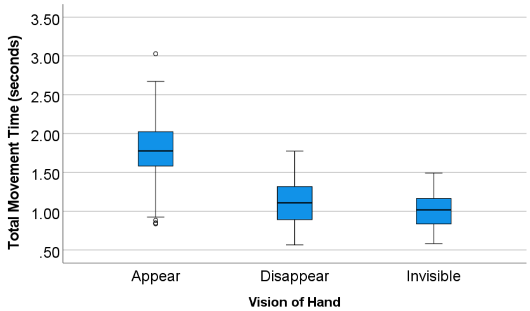

Figure 10: Experiment 2: Mean Total Movement Time (When Inducers Swap in Size) Across Vision of Virtual Hand Conditions.

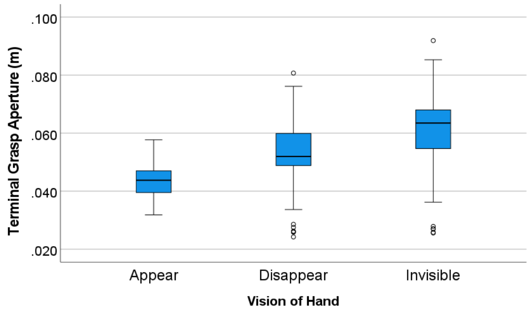

Figure 11: Experiment 2: Mean Terminal Grasp Aperture (When Inducers Swap in Size) Across Vision of Virtual Hand Conditions.

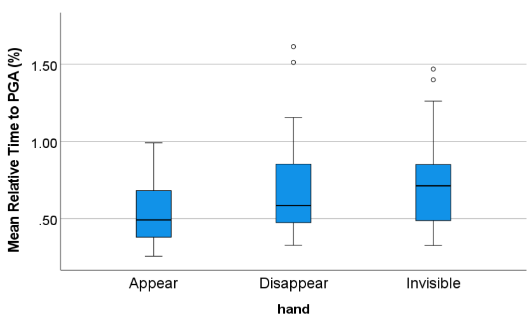

Figure 12: Experiment 2: Mean Relative Time to PGA (When Inducers Swap in Size) Across Vision of Virtual Hand Conditions.

When inducers swapped in size (Figures 10, 11, and 12), vision of virtual hand affected, all kinematic variables were not normally distributed, thus, the nonparametric Friedman test was used to examine the difference among HAND conditions, and the results showed a significant difference for MT ($\chi^2(2) = 54.875$, $p < 0.001$), PGA $\chi^2(2) = 55.167$, $p < 0.001$), Relative Time to PGA ($\chi^2(2) = 21.125$, $p < 0.001$), Relative Time to PRV ($\chi^2(2) = 12.875$, $p = 0.002$), Grasp Plane Error ($\chi^2(2) = 63.500$, $p < 0.001$), and TGA ($\chi^2(2) = 55.792$, $p < 0.001$). As revealed by the multiple pairwise comparison using Nemenyi's procedure, the pattern of results for MT, PGA, Relative Time to PGA, and TGA repeated those seen earlier when inducers appeared or disappeared, in addition, the peak reaching velocity (M=0.73m/s, SD=0.17) occurred earlier in the condition where the hand was visible in the end of movement (M=53%, SD=0.18) than the other two conditions where the hand was invisible in the end of movement (Disappear, M=68%, SD=0.30)(Invisible, M=72%, SD=0.26), suggesting that participants slowed down reaching to focus on adjusting the grip with vision of hand. Moreover, the vision of the hand in the end resulted in the least grasping error in the frontal plane (M=0.01m, SD=0.01), compared to the other two conditions (Disappear, M=0.2m, SD=0.01)(Invisible, M=0.04m, SD=0.02) where the vision of the hand was absent in the end, however, the early vision of the hand still helped to reduce the grasping error in the frontal plane when compared with the condition where the hand is invisible the entire time.

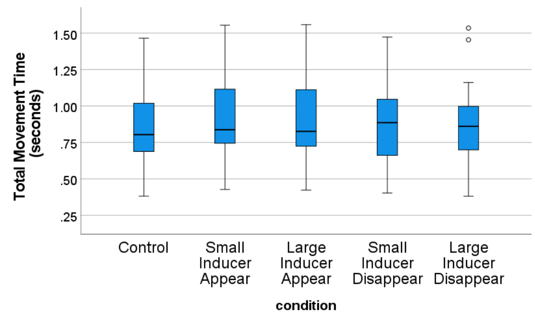

Figure 13: Experiment 3: Mean Total Movement Time Across Conditions where Inducers Appear and Disappear

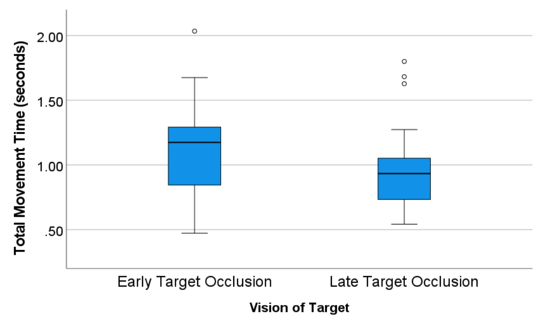

Figure 14: Experiment 3: Mean Total Movement Time Across Early and Late Target Occlusion Conditions

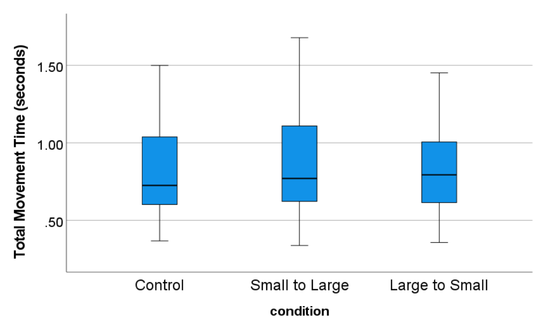

Figure 15: Experiment 3: Mean Total Movement Time Across Conditions where Size of Inducers are Swapped

5.3 Experiment 3

The illusion effect continued to be absent but constrained the grasp aperture with the presence of inducers. However, the vision of contextual objects only affected MT, nothing else. When inducers were constant (Figure 13), MT was normally distributed and examined by rmANOVA, and early occlusion (M=1.09s, SD=0.37), rather than late occlusion (M=0.96s, SD=0.29), of the contextual target resulted in a longer MT ($F_{1,\ 12}$ = 7.48, p = 0.018).

When inducers appeared or disappeared (Figure 14), MT was not normally distributed and examined by Friedman test, and the effect continued to be significant, $\chi^2(1)$ = 9.615, p = 0.002, with early target occlusion taking longer (M=0.91s, SD=0.27) than late target occlusion (M=0.85s, SD=0.26). However, this effect was absent ($\chi^2(1)$ = 0.026, p = 0.873) when inducers swapped in size (Figure 15). Such a finding, combined with the findings from Experiment 1 where vision of contextual object was always visible, suggests that the timing or planning of the movement to interact with the object requires the vision of the object, however, once the movement is initiated, the vision of contextual object would not affect the kinematics and outcomes of movement unless it could be used in combination with the vision of hand.

6 Discussion

Using pictorial 2-D elements, previous studies have shown consistently that Ebbinghaus illusion affects people's judgment on the size of a contextual 2D object in both physical and virtual environment. The current study has clearly demonstrated that Ebbinghaus illusion is robust to affect people's perception on the size of a contextual 3D object in the immersive virtual environment. Compared to the study of Finney & Jones [8], we enhanced the perceptual judgment task by asking participants to create and adjust the target ball size directly in VR while seeing the ball with or without contextual features, allowing for measure of immediate perception in VR. Nevertheless, the illusion effect was found significant, suggesting that human perception is contextual and susceptible to the change of environment, at least in an immersive virtual environment. With regard to the illusion effect on the action used to interact with contextual object, previous studies merely examined the kinematics or outcomes of movement without manipulating the visual information available for the action. This gap is now filled by the current study where we systematically manipulated the vision of contextual features, vision of hand, and vision of contextual object in VR and simultaneously tracked the movement of reaching-to-grasp the object.

A notable finding stemmed from all three experiments was that the Ebbinghaus illusion ceased to affect the action used to interact with the object. The contextual features (large or small inducers) only made the grasp aperture (PGA & TGA) smaller than normal, indicating that they were treated as obstacles to constrain people's opening and closing of fingers during reaching to grasp the object. Such a finding supported the dissociation theory [13] and may suggest that the interactive action demands a focal attention on the target itself independent of its surrounding feature. Although the depth perception induced by Ebbinghaus illusion was not assessed in the perceptual task, its effect on action was seen, however, only in the dynamic illusion condition where the inducers swapped from large to small, people grasped the object closer to the body. This finding conformed to a previous study [8], suggesting that the illusion effect on depth perception was asymmetrical in that the object would be perceived closer when surrounded by smaller inducers, but not farther by larger inducers. Interestingly, when the effect of illusion-induced depth perception was evaluated in action, it can be only seen when the illusion was induced by dynamical change of contextual features involving dramatic reduction in size of inducers. Perhaps, the dramatic change of contextual features yielded the optic flow information that has been shown to affect self-motion and depth perception in a 3D environment [6].

Finally, the current study demonstrated that the vision of virtual hand and the vision of contextual target both played a strong role in visually guided interactions with 3D objects in VR. Specifically, the vision of the contextual target was needed for planning of the action in the beginning, while the vision of virtual hand was needed to fine-tune the action in the end. This finding, though limited to both a virtual environment and object, was consistent with previous studies showing that the sight of hand was unnecessary in early planning and execution of reaching toward a target [15], but was required for final grasping of the object [4]. Especially when the haptic information about the object was unavailable, more time was needed to form the precise grasp aperture by visually detecting the margin between the fingers and the rim of object. In this sense, to improve design of interactive tasks in a haptic-free 3D virtual environment, the vision of hand must be provided near the target object to ensure successful interaction. A potential implementation could resemble using the appearance of the virtual hand when near an interactable object like the change of text color is used when hovering over a link with a mouse to highlight its interactable status. It would serve the usual highlighting purpose but would also confer the benefits discussed above. It is unclear if this design concept could be adapted to work for AR or the real world, as we have explored this phenomenon entirely within a VR setting.

7 Conclusion

In summary, we presented a series of experiments to evaluate the effects of 3D Ebbinghaus illusion with dynamic surrounding features on the task of reaching to grasp a 3D object in an immersive virtual environment. An innovative 3D perceptual judgment task was implemented first to determine if the static illusion affected the perception of the size of the 3D object. This paper presented the results of a series of experiments that investigated the kinematics of reach-to-grasp task through conditions where we experimentally manipulated the visual gain and loss of the 3D contextual inducers, the participant's virtual hand, and the entire 3D contextual object.

In conclusion, the results revealed a significant illusion effect on the perceived size of the 3D object in the judgment task, however in the reach-to-grasp task, only grasp aperture (both peak and terminal) was reduced when the 3D object was surrounded by 3D contextual inducers regardless of their size. A significant undershooting of the object (depth error) was noticed when the contextual inducers changed size from large to small in the midway of reaching toward the object, evidencing the tau effect. We also found that the visual presence of the virtual hand is required for final adjustment of grasp to be accurate, which resulted in increased movement time to complete the task with an Ebbinghaus illusion in effect. The visual presence of the contextual object is required for planning of movement, so the early visual loss of that will result in a longer movement time to complete the task and more time spent on reaching than grasping. The findings from this study are meaningful to the current understanding of how visually guided interactive actions unfolds in a haptic-free virtual environment, and this study can improve 3D user interface design to enhance user experience for immersive virtual environments. Interaction designers can develop novel interaction methods that leverage perceptual gains from this Contextual Illusion to reduce errors and completion time as a result of reducing or eliminating the visual contextual information, such as the virtual hand, to improve interaction performance.

Acknowledgments

The authors would like to thank the participants for their time and patience in completing a series of experiments. The authors would also like to thank Dr. Boyi Dai at University of Wyoming for his consultation to motion analysis and help in providing MATLAB codes for calculating all kinematic variables, which were then transformed and validated in the Unity Engine. This project was funded in part by a University of Wyoming INBRE Graduate Assistantship. As part, this project was supported by a grant from the National Institute of General Medical Sciences (P20GM103432) from the National Institutes of Health. The content does not necessarily represent the official views of the National Institutes of Health.

REFERENCES

[1] S. Aglioti, J. F. DeSouza, and M. A. Goodale. Size-contrast illusions deceive the eye but not the hand. *Current Biology*, 5:679–685, 06 1995. doi: 10.1016/s0960-9822(95)00133-3

[2] S. Alphonsa, B. Dai, T. Benham-Deal, and Q. Zhu. Combined visual illusion effects on the perceived index of difficulty and movement outcomes in discrete and continuous fitts' tapping. *Psychological Research*, 80:55–68, 12 2014. doi: 10.1007/s00426-014-0641-x

[3] A. U. Batmaz, M. D. B. Machuca, D. M. Pham, and W. Stuerzlinger. Do head-mounted display stereo deficiencies affect 3d pointing tasks in ar and vr? *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 12 2019. doi: 10.1109/vr.2019.8797975

[4] C. Bozzacchi, R. Volcic, and F. Domini. Effect of visual and haptic feedback on grasping movements. *Journal of Neurophysiology*, 112:3189–3196, 12 2014. doi: 10.1152/jn.00439.2014

[5] A. J. Bremner, M. J. Doherty, S. Caparos, J. de Fockert, K. J. Linnell, and J. Davidoff. Effects of culture and the urban environment on the development of the ebbinghaus illusion. *Child Development*, 87:962–981, 04 2016. doi: 10.1111/cdev.12511

[6] V. Cornilleau-Pérès and C. Gielen. Interactions between self-motion and depth perception in the processing of optic flow. *Trends in Neurosciences*, 19:196–202, 05 1996. doi: 10.1016/s0166-2236(96)10025-4

[7] M. J. Doherty, N. M. Campbell, H. Tsuji, and W. A. Phillips. The ebbinghaus illusion deceives adults but not young children. *Developmental Science*, 13:714–721, 08 2010. doi: 10.1111/j.1467-7687.2009.00931.x

[8] H. Finney and J. Jones. Asymmetric effects of the ebbinghaus illusion on depth judgments. *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 03 2020. doi: 10.1109/VR46266.2020.00-25

[9] V. Franz, K. Gegenfurtner, H. Bülthoff, and M. Fahle. Grasping visual illusions: No evidence for a dissociation between perception and action. *Psychological Science*, 11:20–25, 01 2000. doi: 10.1111/1467-9280.00209

[10] M. P. Furmanek, L. F. Schettino, M. Yarossi, S. Kirkman, S. V. Adamovich, and E. Tunik. Coordination of reach-to-grasp in physical and haptic-free virtual environments. *Journal of NeuroEngineering and Rehabilitation*, 16, 06 2019. doi: 10.1186/s12984-019-0525-9

[11] S. Glover and P. Dixon. Dynamic effects of the ebbinghaus illusion in grasping: Support for a planning/control model of action. *Perception Psychophysics*, 64:266–278, 02 2002. doi: 10.3758/bf03195791

[12] A. M. Haffenden and M. A. Goodale. The effect of pictorial illusion on prehension and perception. *Journal of Cognitive Neuroscience*, 10:122–136, 01 1998. doi: 10.1162/089892998563824

[13] A. M. Haffenden, K. C. Schiff, and M. A. Goodale. The dissociation between perception and action in the ebbinghaus illusion. *Current Biology*, 11:177–181, 02 2001. doi: 10.1016/s0960-9822(01)00023-9

[14] M. D. B. Machuca and W. Stuerzlinger. The effect of stereo display deficiencies on virtual hand pointing. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 11 2019. doi: 10.1145/3290605.3300437

[15] M. E. McCarty, R. K. Clifton, and R. R. Collard. The beginnings of tool use by infants and toddlers. *Infancy*, 2:233–256, 04 2001. doi: 10.1207/s15327078in0202_8

[16] A. Milner and M. Goodale. Two visual systems re-viewed. *Neuropsychologia*, 46:774–785, 01 2008. doi: 10.1016/j.neuropsychologia.2007.10.005

[17] F. Pavani, I. Boscagli, F. Benvenuti, M. Rabuffetti, and A. Farnè. Are perception and action affected differently by the titchener circles illusion? *Experimental Brain Research*, 127:95–101, 06 1999. doi: 10.1007/s002210050777

[18] P. van Donkelaar. Pointing movements are affected by size-contrast illusions. *Experimental Brain Research*, 125:517–520, 04 1999. doi: 10.1007/s002210050710