Effects of Action Intention, Binocular Depth Cues, Motion Parallax, Haptic Feedback, and Body Posture on the Perception of the Ebbinghaus Visual Illusion

Ruitao Lin

Follow this and additional works at: https://scholarworks.wm.edu/honorstheses

Part of the Cognition and Perception Commons

Recommended Citation
https://scholarworks.wm.edu/honorstheses/1386
Effects of Action Intention, Binocular Depth Cues, Motion Parallax, Haptic Feedback, and Body Posture on the Perception of Ebbinghaus Visual Illusion

A thesis submitted in partial fulfillment of the requirement for the degree of Bachelor of Science in Psychology from The College of William and Mary

by

Ruitao Lin

Accepted for (Honors, High Honors, Highest Honors)

Dr. Peter Vishton, Director

Dr. Christopher Hall

Dr. Mainak Patel

Williamsburg, VA
April 24, 2019
Effects of Action Intention, Binocular Depth Cues, Motion Parallax, Haptic Feedback, and Body Posture on the Perception of the Ebbinghaus Visual Illusion

Ruitao Lin
College of William & Mary
Williamsburg, VA, USA
Abstract

Researchers have long observed different illusion magnitudes in verbal response tasks and visually-directed action tasks. The cause of such differences has been the topic of debate. The “two visual systems hypothesis” (TVSH) suggests that two separate visual pathways independently control a certain type of tasks. According to this theory, the difference in illusion magnitudes is caused by the different performance of these two pathways. An alternative theory is the “two modes of processing” (TMOP) hypothesis, which states that the two visual processing modes function within a single visual pathway but weigh the same set of visual information differently. According to this theory, the drop of illusion magnitudes in visually-directed action tasks is the result of such different weights. The three experiments presented here focus on the effect of motion parallax and binocular depth cues, haptic feedback from 3D target disks, and body postures, respectively. Results suggest that while haptic feedback and body postures are critical to the reduction in illusion magnitudes, motion parallax and binocular depth cues seem to be irrelevant. Limitations and future directions are suggested.

Keywords: Ebbinghaus Illusion, motion parallax, binocular depth cues, haptic information, body postures, embodied cognition
Effects of Action Intention, Binocular Depth Cues, Motion Parallax, Haptic Feedback, and Body Posture on the Perception of the Ebbinghaus Visual Illusion

While human visual perception is very accurate, it is not always perfect, especially in the cases of size and distance perception. Contextual cues and background information can cause errors in people’s perception of object size and location. In the past century, scientists have been trying to investigate and characterize these perceptual errors and found that their magnitudes vary under different conditions.

For instance, in the case of the Ebbinghaus illusion, the measure of illusion magnitude is an observer’s point of subjective equality (PSE). The Ebbinghaus illusion occurs when two disks, identical in size, are surrounded by arrays of disks of different size (Figure 1). The central disk surrounded by the small-disk array (the standard disk) is perceptually larger than the central disk inside the large-disk array (comparison disk). When the size of the comparison disk increases gradually, at a certain point its size will appear to be equal to the size of the standard disk. The sizes of the perceptually equal, but physically different comparison disk define the observer’s PSE. The relative sized of these two disks when they were perceptually equal is often used to indicate the magnitude of an illusion for a certain observer. For instance, if a 30 mm diameter disk is perceived to be the same as a 28 mm disk, the magnitude can be calculated as

\[
\frac{(30-28)}{28} \times 100 = 7.1\%.
\]

The magnitudes of the illusions seem to vary when different response tasks are measured. When visually-guided or visually-directed actions, like pointing at or grasping the object, are assessed, perception is often more accurate than when people respond with verbal judgement (Aglioti, DeSouza & Goodale, 1995; Haffenden & Goodale, 1998; Bridgeman, Kirch & Sperling, 1981; Wong & Mack, 1981; Loomis, Da Silva, Fujita, & Fukusima, 1992).
Researchers have studied many different types of perception tasks. These tasks can be organized into three categories: visually-guided action, visually-directed action, and verbal response. Visually-guided action is defined as movement guided by continuously available visual information. For example, how a person sees and picks up a water bottle in front of him or her is a visually-guided action. Visually-directed action is similar, but the visual information is not directly accessible but stored in memory (Loomis et al., 1992). Visually-directed action can be illustrated as the person performs a reaching action after the water bottle has been removed, thus eliminating the visual stimulus. Another type of behavioral response that interests researchers is the verbal response judgement. Indeed the vast majority of research on visual perception has focused on this type of task. It is closely associated with our intuitive sense of “conscious” perception. In the example of the water bottle, verbal response judgement would be how the person verbally describes the size or position of the bottle without performing any visually-directed or guided action.

Many scientists have believed that visual perception of an object generates a singular mental representation of the object in the observer's brain. Such a mental representation stores the size and distance information of the object. Therefore, for the same physical stimulus, an observer should exhibit identical magnitudes of any size illusions across all types of response tasks. Loomis et al. (1992), however, found that people judge the same size information in their action space differently when engaged in different tasks. Action space is defined as the space of an individual’s public action (Cutting & Vishton, 1995). Its limit is beyond arm reach but within 30 meters. In the Loomis et al. visual matching task, participants saw two targets displayed horizontally in front of them and estimated the interval between the targets. The distance between a participant and targets never exceeded 12 m, so that the targets were always within the participant’s action space. They then asked researchers to adjust two target rods in depth so that the distance between the target rods matched the
estimated horizontal interval (Figure 2). In general, participants’ estimations were larger than the true size of the interval. When the distance between the participant and the target interval increased, the amount of overestimation increased. The data suggest that participants experience a “depth compression illusion.” For each additional meter in distance, even with constant interval size, participants perceive progressively smaller amounts of distance. The magnitude of this compression illusion is reduced when the target is near the person.

Loomis et al. (1992) also assessed a walking-based measure of distance perception. In the walking task, participants viewed the same interval. They then walked with closed eyes until they thought their walking distance matched the interval. Essentially participants closed their eyes and attempted to walk to the target. Since they were not able to see during walking, participants were not able to make any correction from visual feedback. Their walking distance depended solely on their mental representation of the interval. If they use the same mental representation of distance for all tasks, the illusion in the estimation task would also be present in the walking task. The participants should have walked past the targets.

Although participants still overestimated the size of intervals, the errors were much smaller than in the perceptual matching task. The error was also independent of the distance between participants and the target. If the size judgement in different response tasks relies on a singular mental representation generated by a single visual processing stream, then there should still be a large effect of the illusion and a compression effect in the walking task. The difference in the magnitude of illusion reported in Loomis et al. (1992) suggests that there are two separate processing mechanisms in the visual system. One mechanism is responsible for visually-directed action and the other is responsible for verbal response judgement of visual information.

Other pictorial illusions also seem to influence our conscious perception and our action differently. Bridgeman, Kirch, and Sperling (1981) demonstrated such a difference in
illusion perception through a moving frame paradigm. In their experiments, participants were required to focus on a target dot. The background pattern moved abruptly while the target dot remained fixed (Figure 3). Then the whole display disappeared. In the verbal response task, the participants saw a few candidate dot locations. They verbally stated which one was the true location of the dot. In the action task, they were instructed to point at the dot position. The motion of the background induced a perceptual displacement of the target dot to a different location. For instance, if the frame moved left, participants would later remember the dot as positioned slightly to the right of its actual position. Although participants verbally reported the induced perceptual change, they could always point at the actual, unchanged location of the target after the target and the background disappeared. In this experiment, visually-directed action seemed to be less affected by the illusion than verbal response about the target position. This result is consistent with the idea that two separate visual processing mechanisms govern visually-directed action and verbal response judgement.

Eye movement data have also provided evidence of different response performance across different types of task. The study of Wong and Mack (1981) used a procedure similar to that of Bridgeman et al. (1981), but participants were asked to look at the dot position in the action task. Participants reported the induced illusion in the verbal response task. Eye tracking, however, revealed that participants consistently directed their fixations to the true location of the target dot instead of the perceived position. While the moving frame affected the participants’ verbal response behaviors, it seemed to have little impact on their eye movements, again suggesting the existence of two different visual processing systems.

If visually-directed action, visually-guided action, and verbal response judgement were all based on a single mental representation, then these tasks would produce illusions in size, distance, and position perception with identical magnitudes. Results from previous studies, however, have provided contradictory evidence. There seem to be separate
mechanisms controlling verbal response and object-directed actions, resulting in different magnitudes of illusions in the two types of tasks.

The cause behind such difference in performance has been the topic of many research projects. Some scientists argue for the Two Visual Systems Hypothesis (TVSH), which suggests that the different performances are caused by two separate visual systems in the brain (Milner & Goodale, 1995). According to this class of theories, a cognitive system controls visual analysis and identification and activates the “what” or “perception” visual pathway. A separate sensorimotor system controls visually-guided actions, activating the “where” or “action” visual pathway (Tseng & Bridgeman, 2011). Other scientists have countered this hypothesis through experiments. They state that, instead of two distinct visual pathways, there may be two modes of computation within a common set of neural circuits. Vishton and colleagues have described this as the Two Modes of Processing hypothesis (TMOP) (Vishton, Stephens, Nelson, Morra, Brunick, & Stevens, 2007). The following sections will discuss the TVSH and the TMOP in detail.

Two Visual Systems Hypothesis

In the previous section, I introduced studies that illustrate the different perception performance in verbal response tasks and visually-directed or guided action tasks. These studies suggest the existence of separate mechanisms in human visual processing. The cause of different performance and the nature of the separate mechanisms have been the center of many debates. Some scientists have argued that the results suggest the TVSH.

Evidence for the TVSH has been found in research projects involving brain injured patients (Perenin & Vighetto, 1988; Jakobson, Archibald, Carey, & Goodale, 1991; Goodale, Milner, Jakobson, & Carey, 1991). Patients with damaged posterior parietal cortex often have trouble controlling hand movements, while patients with damaged occipital lobe often show deficits in object recognition.
Optic ataxia is a deficit of coordination and accuracy of visually-guided movements without perceptual, visuomotor, and attention disorder. Such a disorder, caused by lesions of the posterior parietal lobe, often produces a disturbance of reaching and grasping movements (Perenin & Vighetto, 1988; Jakobson et al., 1991). In their study of optic ataxia patients, Perenin and Vighetto recorded patients’ response and arm movement during three tasks. In the visual space perception task, patients verbally identified a dot position and line orientation. In the object reaching task, patients reached out to grasp an object in front of them. In the hand orientation task, patients fitted their hand into a tilted gap in front of them. Results showed that these patients were incapable of controlling the grasping motions and orientation of their hands, even though they were able to perceive the orientation of the slot. An individual kinematic analysis of patient V.K. showed that the patient had directional misreach in pointing and grasping tasks as well as significantly different movement trajectories when compared with typical participants (Jakobson et al., 1991).

Patient D.F., who had damage in the lateral occipital lobe and parasagittal occipitoparietal region, showed different performance when compared to optic ataxia patients (Goodale et al., 1991). She was incapable of verbally stating object orientation, but she successfully controlled the orientation of her hands to grasp an object she could not perceive.

Researchers have also used maximum grip aperture (MGA) as a measurement of size perception. When an individual picks up an object, his or her thumb and index finger open up during the movement then close to grasp the object. The aperture between the thumb and index finger typically reaches its maximum approximately halfway through the reaching movement. MGA is highly correlated with the object size, therefore it is used as an assessment of how visual size perception is reflected in the action tasks (Paulignan, MacKenzie, Marteniuk, & Jeannerod, 1997). It is measured through sensors attached to the participant’s thumb and index finger (Figure 1). Regarding the MGA of patient D.F., her
performance was not different from that of the control subjects. Results from these experiments and case studies suggest that visual perception and visually-guided actions are controlled by different neural substrates.

Supporting evidence for TVSH is also present among people without brain injury. Scientists have demonstrated the dissociation between two visual pathways through the perception of Ebbinghaus illusion (Aglioti et al., 1995; Milner & Goodale, 1995; Haffenden & Goodale, 1998).

When a person reaches out to pick up an item, the maximum grip aperture (MGA) (distance between the thumb and index finger) is highly correlated with the actual size of the object (Paulignan et al., 1997). As a result, MGA can be used as an indicator of the visual size registration that is used to control grasping actions. Aglioti and colleagues presented 3D Ebbinghaus illusion displays to participants and asked the participants to make judgements about the sizes of the central disks. When the two central disks appeared to be the same size, participants were asked to pick up the left disk. If the two disks were perceptually different in size, participants were required to pick up the one on the right. For experimental control the choice was reversed for half of the participants. Each participant was tested individually to establish a pair of disks that were perceptually equivalent in size (physically different) to him or her. The researchers tracked the onset time of the grasping action, the trajectory of the movement, and the maximum grip aperture through infrared light emitting diodes.

Results showed that when disks were perceived to be of the same size, the mean onset time was longer than when disks were perceptually different in size. Grasping aperture was significantly affected by the illusion but the effect was significantly smaller than what researchers found with the size judgement tasks. When two disks were perceived as the same size while physically different, the average grasping aperture was larger for the larger disk
than for the smaller disk. This result indicates that the calibration of grasping is to certain extent less affected by visual illusion than is conscious perception.

Despite this result, the visual illusion still influenced MGA. In cases where two identical disks were perceived to be of different sizes, average MGA was larger for the disk surrounded by small circles (which was judged as the larger disk) than for the disk surrounded by large circles. The effect of the illusion on grasping aperture is significantly smaller and less variable than the effect of the illusion on visual perception.

A modified study also provided similar results and offered support for the TVSH theory (Haffenden & Goodale, 1998). In the previous study of Aglioti et al. (1995), participants may adjust their grip aperture after reaching out based on visual feedback. As a result, the maximum grip aperture would not be an accurate measurement of initial scaling. To avoid visual feedback, Haffenden and Goodale conducted their experiments under “open-loop” conditions. Open-loop condition is defined as the condition in which a participant is not able to see how they perform the action. On the contrary, a closed-loop condition is where the participant’s hand movement is visible to him or her. An overhead light was shut off during reaching movements so that participants were not able to see their hands. Researchers also changed the verbal judgement procedure into manual estimation. Participants were instructed to use the distance between their thumb and index finger as an estimation of the size of the disk. Verbal judgement tasks in previous research were dichotomous (either of equal size or different in size). The change of tasks ensured a continuous measurement of both visual perception and grasping. Haffenden and Goodale also introduced two additional backgrounds, a blank background and an equal-disk background, in which the surrounding arrays were consisted of circles of equal size. The new backgrounds were designed to examine the effect of surrounding circles on manual estimation and grip aperture.
With illusion-inducing backgrounds, when the two center disks were physically different but perceptually identical, participants’ grip apertures were scaled to the physical sizes of the disks. Their manual estimations, however, were the same for both the larger and the smaller disks, meaning that the estimation apertures were scaled to the perceptual sizes. When two identical disks were presented, grip apertures were resistant to illusion. Manual estimation of disks in small circle arrays was larger than the estimation of disks in large circle arrays, indicating the effect of illusion on visual perception. In blank background and equal-disk backgrounds, manual estimation and grip aperture were both scaled to the actual sizes of the disks presented to the participant. The equal-disk background produced smaller MGA than blank background did. In contrast, manual estimation was larger under equal-disk background than under blank background.

Studies on Ebbinghaus illusion and the research on brain-injured patients confirm the performance difference between conscious visual perception and visually-guided actions. These results provide evidence for the TVSH theory: one stream controlling object perception and another controlling visually-guided action. Although such a model can successfully explain the performance under different tasks, scientists have also found contradictory evidence, suggesting the model is inaccurate.

Shortcomings of TVSH

The result of Aglioti et al. (1995) and the TVSH has faced many criticisms (Franz, Gegenfurtner, Bulthoff, & Fahle, 2000; Franz & Gegenfurtner, 2008; Smeets, Brenner, de Grave, & Cuijpers, 2002). Franz and colleagues (2000) criticized the methodology of the original study. They argued that the dissociation was observed due to confounding differences between the perception and action procedures. When they carefully adjusted the experiment methods to eliminate these confounds, they observed no evidence for the dissociation.
Franz et al. (2000) argued that in the original experiments by Aglioti et al. (1995), the effect of the illusion cannot be directly compared because the two tasks involved asymmetrical measures. Franz et al. conducted three modified experiments. In the first experiment, the original verbal judgement task was replaced by a different version of the perception task. Participants were asked to manually adjust the size of a separate comparison disk to match the size of one of the central disks in the illusion display. The final size of the separate comparison disk was recorded as the perceived size. The grasping task remained the same. This experiment was conducted under open-loop conditions in which the participants were not able to see the display and their hands during grasping. MGA and grasp trajectories were recorded.

A significant effect of the illusion was present both in the perception task and in the grasping task. Both the perceived size and MGA were linearly related to the actual size of the central disk, so researchers could compare the effect of the illusion directly. The effect, however, was the same for the two tasks, contradictory to the original findings. Also, participants who showed larger effects of the illusion in the perception task also showed larger illusion effects in the grasp task. The perceptual illusion in the study by Aglioti et al. (1995) was larger than the perceptual illusion reported by Franz et al. (2000), while the grasping illusion stayed the same. Franz and colleagues hypothesized that the enhancement was due to the nature of the perceptual task in the original study. They conducted a second experiment to test their hypothesis.

In their second experiment, Franz and colleagues (2000) conducted single context comparisons, separate comparisons, and the direct comparisons. In the single context comparison, a separate, isolated disk was used to compare with only one of the illusion displays (only the small-disk array or only the large-disk array) at a time. The effect of the large-disk array and the effect of small-disk array were added to determine illusion
magnitude. In the separate comparison situation, a separate comparison disk was present. Participants adjusted the size of that comparison disk to match the size of one of the central disks.

Single context comparisons and separate comparisons showed similar magnitudes of illusion. The effect of the illusion in the direct comparison tasks was larger than the sum of the effects in the two single context comparisons and the sum of effects in the separate comparisons. In the original study of Aglioti et al. (1995), the direct comparison task produced a larger illusion magnitude than viewing the two displays separately. As a result, the experiment design in the original studies did not perfectly match the two tasks and the conclusion of dissociation is not valid.

In their third experiment, Franz and colleagues (2000) further investigated the procedure in the original study. Participants first made direct comparisons, then performed separate comparisons. The illusion effect that resulted from direct comparisons was significantly larger than the sum of illusion under separate comparisons. Since participants only reached for one part of the display in the action task, the smaller separate comparison component of the illusion was perhaps to be expected. In general, the perceptual task and the grasp task generated similar illusion magnitudes when matched properly. The larger effect of perceptual task in the original study, according to Franz et al., was due to a failure of additivity.

As mentioned above, Franz et al. (2000) found no performance difference in the visual perception and the visually-guided action tasks, which contradicts the prediction of the TVSH. Franz and Gegenfurtner (2008) proposed three possible explanations to reconcile this discordance. The redundant illusion hypothesis suggests that the Ebbinghaus illusion is generated in the two visual systems separately and have the same magnitude in both systems. A second possibility is that there is strong communication between the two systems, which
makes the effect of illusion equal in both systems. The third hypothesis is that the Ebbinghaus
illusion is created in the brain before the two visual systems separate from each other.

Although these three hypotheses successfully justified the result of Franz et al. (2000)
under the premise that two visual systems exist, they nevertheless undermine the importance
of the TVSH theory. The existence of the two separate systems would be redundant and
inefficient in utilizing available resources (Franz & Gegenfurtner, 2008). Other scientists
have suggested different theories. For example, Vishton et al. (2007) have developed the Two
Modes of Processing (TMOP) hypothesis. They argue that the brain processes the same set of
visual information differently when participants engage in visual perception and visually-
guided action tasks.

Two Modes of Processing Hypothesis

Vishton et al. conducted three experiments to test the TMOP hypothesis (2007). In the
first experiment, they replicated the previous findings of Aglioti et al. (1995). In the setup,
the standard disk inside the small circles array was 28 mm in diameter. In each trial, the
standard disk was paired with one of six other disks, with diameters ranging from 27 mm to
33 mm. Half of the participants completed 24 verbal trials in which they verbally stated
which disk appeared larger. The other half of participants completed 12 verbal trials and 12
grap trials, in which they picked up the disk that looked larger. Both plain backgrounds and
illusion backgrounds were presented. The order of the background conditions and the order of
the comparison disks were randomized. Researchers recorded the MGA for participants who
completed the grasp trials. Data analysis showed that the grasping task in general generated
40% less illusion than the verbal judgement task did, which is consistent with previous
studies. A positive correlation between illusion magnitude and grasp/verbal choice suggest
that the response was not produced by independent visual pathways. Even when viewing
identical displays and making the same perceptual choice, participants who were preparing to reach showed a smaller magnitude of the illusion.

The second experiment ruled out the possible confounding variable of haptic feedback in the first experiment. Instead of picking up the disk, these participants touched the top of the disk they perceived as larger rather than grasping it. Half of the participants received instructions of the action task after they finished verbal trials, while the other half received complete instruction of both the verbal task and the action task at the beginning of the experiment. The purpose of this condition was to test whether the intention to touch could reduce the effect of illusion.

Participants who received the target touch instructions after the verbal trials produced data similar to that of the participants in the first experiment (Vishton et al., 2007). The effect of the illusion was reduced if participants were asked to touch the disk. Participants who received complete instructions at the beginning of the experiment not only showed reduction of illusion in the touch tasks, but also showed smaller effects of the illusion for both verbal trials and touch trials compared to the other half of participants. Such results support the idea of the interaction between intention to touch and visual perception.

In general, the results of Vishton et al. (2007) support the idea that the intention to reach influences visual processing. While these researchers successfully replicated the result reported by Aglioti et al. (1995), the reduction of illusion magnitude was also observed in the verbal response tasks when participants were given complete instruction or performed action tasks prior to verbal response tasks. Such results suggest that there is only one visual processing pathway with different modes of computation.

The experiments conducted by Vishton et al. (2007) show that touching, instead of grasping, is sufficient in generating the reduced effect of Ebbinghaus illusion. What’s more, when participants knew that they would act upon the disk, the magnitude of illusion also
decreased. Vishton et al. interpret their findings as evidence that the intention to act upon the disk can significantly reduce the effect of Ebbinghaus illusion. The similar illusion magnitude in verbal judgement and touching choice (for verbal task: 6.18%; for touch task: 5.54%) with pre-instruction suggest that a single visual system, rather than two separate systems, is responsible for both verbal judgement and visually-guided action tasks. Vishton et al. argue that action planning changes how the system processes visual information.

The TMOP hypothesis suggested by Vishton et al. is also consistent with the finding of Witt, Proffitt, and Epstein (2005). Witt and colleagues investigated whether people’s ability to reach with a tool would influence their perception of distance through three experiments. In the first experiment, researchers tested whether holding a tool would alter a person’s perception of distance. Each participant completed two blocks of trials. In one block of trials they held a baton and in the other block they did not have any tool. A stimulus circle was briefly flashed on to a table surface. Participants would either touch the location of the circle or estimate the distance. When participants held the baton, their estimation of the location was closer than when they reached with finger, suggesting that they perceived the distance differently when holding the baton.

In a second experiment, Witt et al. (2005) further investigated the influence of reachability through a matching task. Participants were instructed to adjust the distance between two comparison circles to match their perceived distance of a projected target. After the estimation, the target disappeared and participants reached to where the target was, either with finger or with the baton. Similar to the result in their first experiment, when participants reached with their finger, they estimated the distance to be farther than when they reached with the baton. Witt et al. argue that holding the baton expanded “personal near space” such that the object was perceived as being closer.
In the third experiment, participants made estimations but did not reach to the target afterwards. In this case, personal near space did not expand when participants had no intention to reach. Witt et al. (2005) concluded that only the intention to reach to the target affects distance perception. This result is similar to the result of Vishton et al. (2007), in which the intention to grasp altered conscious perception of target size.

A study by Iriki, Tanaka, and Iwamura (1996) suggested a neural mechanism for these performance differences due to action intention. According to their research on macaque monkeys, the monkey’s “reachability” neurons, which fire when an object is within reach distance, adapted to the change in distance when the object was moved further but still within the monkey's reach with a rake.

Neglect patients have also been found to show different behaviors between an object within near space and an object beyond near space. Normal people showed different tactile perception when a light was within reach than when the light was far away. These results suggest that near space is remapped when a tool is used and human visual system is sensitive to the change of reachability. What’s more, these results indicate that an object that is initially beyond reach would appear to be closer when a tool is involved without the change in distance. Other research has shown that human body and effort of action provide metrics for perception.

Research evidence mentioned above has suggested that there is only one visual system with two modes processing. One mode is responsible for facing verbal response judgement and the other is responsible for tasks involving object-related actions. The two distinct modes weigh various size and distance information differently during information processing, and as a result induce different magnitudes of illusions.

Previous research has identified distinct performance in perception when people make verbal judgements and when people perform object-related actions. Some researchers have
suggested the existence of two separate visual pathways and proposed the TVSH theory. Each pathway controls human visual perception in different types of tasks. The TVSH theory is supported by observational data of the behaviors of brain-injured patients, as well as experimental data of perception tasks in people without brain injury. Such a hypothesis, however, has faced many challenges. Some researchers have questioned the existence of two separate visual systems. Vishton et al. (2007) instead proposed the TMOP theory, arguing for two separate processing modes within a single visual system. The observed different perception performance is induced by different weights of visual cues in the two computational modes.

Overview of the Experiments

The three experiments presented in this paper aim to test the cause behind different perception response in different tasks. All three experiments used the Ebbinghaus illusion. Experiment 1 tested the TMOP hypothesis by identifying the visual cues that are weighted differently in the two processing modes. As suggested by Cutting and Vishton (1995), the two most effective visual cues in size perception are binocular depth cues and motion parallax. If these two sources of visual information are removed from the perception process, there would be an increased magnitude of illusion in the action task. Experiment 2 investigated the effect of haptic feedback and 3D nature of the target disks. The illusion display was presented in 2D format on a computer monitor to eliminate haptic feedback. Experiment 3 examined the influence of body posture on illusion perception. In grasp tasks, participants always move their hands to touch or pick up the target disk, placing their hands in the vicinity of the illusion display. Previous research has suggested that hand position alters how people perceive size and allocate attention. The purpose of Experiment 3 is to test whether hand position affects the perception of Ebbinghaus visual illusion and offer a new perspective to the debate between TVSH and TMOP.
Experiment 1: The Effect of Binocular Depth Cues and Motion Parallax on the Ebbinghaus Illusion

As mentioned above, the TMOP hypothesis argues that the two processing modes within a single visual system are responsible for the different magnitudes of Ebbinghaus illusion. When perceiving the size and distance of an object, people utilize a variety of visual information. As the TMOP proposes, the two processing modes have different relative sensitivities to those different visual information. If certain visual depth cues have more weight under the “visually-direct action” processing mode, the magnitude of the Ebbinghaus illusion might differ from that in the “verbal response” processing mode as a result.

Common visual cues that people employ in size and depth perception include binocular disparity, motion parallax, relative size, and occlusion. Among these visual information, I aim to test two factors that work the best when the object is in personal space: binocular disparity and motion parallax (Cutting & Vishton, 1995). In most of the previous studies, the illusion display was presented within a participant’s arm reach so that the participant could perform the grasp task. The space within arm reach of an individual (approximately within 1 m in radius) is defined by Cutting and Vishton as the personal space.

Binocular disparity is defined as the different retinal image of the same object in two eyes. When people view with both eyes an object that is not at the fixation point, the object is projected to different locations on the left and right retina (Figure 4). The difference in angle, position, and size of the two images all provide information to binocular disparity.

Another effective visual depth cue is motion parallax. When two objects, with fixed positions relative to each other, move at the same speed in front of the perceiver, their retinal location also move, but at different speeds. The object closer to the observer moves at faster magnitude. The object further from the observer moves slower. Such a difference allows the perceiver to triangulate the position of the two objects.
Since binocular depth cues and motion parallax are two sources of visual information that work best in the personal space, I aim to test whether they are weighted differently in the two modes of visual processing. If so, then preventing observers from using these cues would result in a similar performance in verbal response judgement and visually-guided action judgement. When perceiving the illusion display, observers tend to move their head. Even a slight head movement can induce sufficient motion parallax information. In order to eliminate their access to motion parallax cues, I used a chin rest to fix the participant’s head position. Regarding binocular disparity, some participants were instructed to wear an eyepatch that blocked the vision of their non-dominant eye.

**Methods**

*Participants.* Two hundred and twenty six College of William & Mary students (132 female, 94 male) participated in the experiment for course credit in PSYC 201/202 or as volunteers. Among these participants, twenty three were excluded because of experimental errors. Nine participants were excluded because they had difficulty performing the grasp task. A total of one hundred and ninety-four participants (113 female, 81 male) were included in the final data analysis.

*Displays and Apparatus.* Participants sat in a chair (45 cm tall) in front of a table surface (73 cm tall X 152 cm wide x 76 cm in depth). For two-thirds of the participants, a chin rest, approximately 30 cm high, was placed by the edge of the table (Figure 5). The height of the chinrest was adjusted for each participant to sit comfortably in upright position. The eyepatch was a pair of glasses with one side blocked by opaque tapes. Two eyepatches were used, one for blocking the participant’s vision from the left eye and the other for blocking the participant’s vision from the right eye. Only one eyepatch was given to each participant based on his or her non-dominant eye.
In the plain background condition, targets were presented on a blank paper (28 cm wide x 21.6 cm in depth). The paper was centered in front of the participant, 21.3 cm from the front edge of the table. The position of the paper was marked by colored tapes. Two marks (one 44 mm from the left edge, one 90 mm from the right edge, and both 97 mm from the front edge) on the paper ensured the consistent placement of target disks.

In the illusion background condition, the target disks were presented on a separate piece of paper. Two marks on the paper, with position identical to those in the plain ground condition, ensured the consistent location of disks between trials. A small-disk array, consisted of 11 disks with 10 mm in diameter, surrounded the mark that is 44 mm to the left edge of the paper. The distance between the closest edge of the disks and the mark is 19 mm. The other mark was surrounded by a large-disk array, which contains five 58 mm black disks. The distance between the closest edge of the disks and the mark is 25 mm.

Target disks were made of black plastic (3 mm thick). The standard disk remained as a 28 mm disk throughout the experiment. The diameter of comparison disks were 27, 28, 29, 30, 31, and 33 mm. A wooden panel frame (30 cm x 35 cm x 2 cm) was used to block participant’s view of the display between trials.

**Design.** I randomly assigned each participant, without replacement, to one of the three viewing conditions (Eyepatch and chinrest, no eyepatch with chinrest, and freeview). In the freeview condition, there was no chinrest nor a glasses with occluding eyepatch. In the eye patch with chinrest condition, participants wore a pair of occluding glasses that blocked their view of the non-dominant eye, and sit their chin on the chinrest. In the no eyepatch with chinrest condition, participants sit on a chinrest without the view-occluding eyepatch. Each participant was randomly assigned to one of two side conditions (the small-disk array on left vs. right). Each participant was randomly assigned to one of two choice condition (picking
smaller disk vs. larger disk). The viewing condition and orientation of display remained throughout the experiment.

In 1 block of trials, each participants viewed 6 sets of disk comparisons (27, 28, 29, 30, 31, 33 mm) in each of 4 background conditions (2 plain vs. 2 illusion). The order of comparison pairs and the order of background was randomized across participants. All participants first experienced four blocks of verbal trials (a total of 24 trials), in which they verbally indicate their selection of disk, followed by four blocks of action response trials (a total of 24 trials). In the action response block, participants indicated their selection using grasp response and were prohibited to use verbal indication.

Procedure. After providing informed consent, participants were tested of their dominant eye and dominant hand. The experimenter then instructed the participant to put his or her chin on the chin rest based on the condition he or she was assigned to. Participants assigned to wearing eyepatch were given the pair of glasses that occluded the vision of their non-dominant eye. Then participants were told that they would be making judgement of relative size of two disks. They should indicate their selection by verbally stating whether the larger (smaller) disk was on their left or right side. After the participant indicated his or her selection, the experimenter put the wood board between the participant and the display while placing new comparison pairs.

After 24 trials, participants in the GRASP condition were instructed to continue making the same size judgement. Instead of verbally indicating their choice, they should reach out and briefly lift up the disk with their thumb and index finger. Participants in the IMAGINE condition were instructed to mentally perform the grasping task without making actual movement. Then they should verbally indicate which disk they selected to pick up.

The entire procedure lasted approximately 20 min.
Data Analysis. For each set of 12 trials conducted under the same background and task condition, I identified the largest comparison disk selected at least once as being smaller than the standard 28 mm disk. For example, in plain backgrounds, regardless of task types, most participants would pick the 27 mm or 28 mm comparison disk as smaller than the standard 28 mm disk. In illusion backgrounds, however, because of the relative sizes of surrounding arrays, people sometimes pick the 30 mm or 31 mm comparison disks as smaller than the standard disk. The size of the largest of these disks was recorded as the criterion value. Effects of condition and response task were assessed using mixed model repeated measures ANOVA. In most cases the sphericity assumption was violated. For all reported results, the Greenhouse-Geisser correction has been applied.

Results and Discussion

Participants’ selection of disk was a function of comparison size in all backgrounds, tasks, and viewing conditions (Figure 6, Table 1). In general the Ebbinghaus illusion produced a significant increase in criterion value, $F(1, 191) = 976.3, p < 0.0005$. Across all three viewing conditions (eyepatch and chinrest, chinrest without eyepatch, and freeview), there was a significant interaction between the effect of backgrounds (plain vs. illusion) and response tasks (verbal vs. grasp), $F(1, 191) = 23.5, p < 0.0005$. The effect of illusion on verbal response trials was 10.3%, while the effect of illusion on grasp tasks was 8.4%. In general the finding was consistent with previously observed data of Vishton et al. (2007), though the drop of illusion magnitude was smaller than reported in the previous study.

There was no significant three-way interaction between viewing conditions (eyepatch and chinrest, chinrest without eyepatch, and freeview), tasks (verbal response vs. grasp), and backgrounds (plain vs. illusion), $p > 0.5$. This result suggests that eliminating motion parallax and binocular depth cues does not affect the reduction of illusion in visually-guided action
tasks. Motion parallax and binocular depth cues seem to have the same weight in the two processing modes.

In the perception of the Ebbinghaus illusion, the magnitude of illusion decreases when people engage in visually-guided action tasks when compared with verbal response tasks. As the TMOP hypothesis proposes, such a change in illusion magnitude is due to different modes of visual processing that are in charge of different types of tasks. The two processing modes have weights of various visual cues, thus the illusion magnitudes in the visually-guided action tasks and the verbal response tasks are different.

To test whether the shift of illusion magnitude is due to the increase weighting of motion parallax and binocular depth information, I eliminated these two visual cues from the participant’s perception process. Motion parallax was removed by applying a chinrest, and binocular depth information was removed through eyepatch.

Participants in the three viewing conditions showed similar reduction of illusion in the grasp task, indicating that the weight of motion parallax and binocular depth cues are equivalent in both processing modes. The reduction of illusion magnitude observed in visually-directed grasping tasks could be attributed to other visual or sensation information. Specifically, in the grasping tasks participants could gain haptic feedback from picking up disks in the preceding trials. In the second experiment I tested whether the observed shift in illusion magnitude is due to the haptic feedback of the physical disks.

Experiment 2: The Effect of Haptic Feedback on Ebbinghaus Illusion Perception

As the results of Experiment 1 suggest, motion parallax and binocular depth cues do not have different weights in the two visual processing modes. When these visual cued were inaccessible, participants still showed similar drop of illusion magnitude in visually-guided action tasks. These two visual cues, however, are not the only information utilized when
people perceive the size of an object. Other size information generated by the different characteristics of the two tasks could contribute to the different illusion magnitudes.

In the visually-guided grasp tasks, participants were instructed to pick up the disk of their selection. The grasp action could provide sufficient haptic feedback of the true size of the disk, which can be used to calibrate the size perception of disks. Such haptic feedback was not available in the verbal response tasks as participants only verbally indicated their selection. Perhaps the sensation feedback provided information to the visual processing mode that controls the size perception in visually-guided action tasks. People actively calibrated their perception based on the feedback from preceding trials. Without the feedback information provided by the grasping action, the visual processing mode controlling verbal judgement tasks cannot calibrate based on previous trials, and thus was more susceptible to visual illusion.

Experiment 2 aims to examine whether the shift in illusion magnitude observed in visually-guided action tasks is generated by haptic information. If so, then the verbal response tasks and visually-guided action tasks would show similar magnitudes of illusion when the targets are impossible to be picked up. In this experiment, the target disks were shown on a computer screen so participants would not acquire haptic feedback of the size of the disks. What’s more, Vishton et al. (2007) suggest that merely touching targets disks can induce the reduction of illusion magnitudes. Experiment 2 also tests whether touching ungraspable disks changes the illusion magnitude.

Methods

Participants. Forty-five College of William & Mary students (27 female, 18 male) participated in the experiment for course credit in PSYC 201/202 or as volunteers. Among these participants, five were excluded because of experimental errors. A total of forty participants (23 female, 17 male) were included in the final data analysis.
Displays and Apparatus. Displays and apparatus were the same as in Experiment 1 except noted here. Both the plain backgrounds and the illusion backgrounds were projected on a computer monitor (Samsung S27E310, 64.3 cm x 46.2 cm x 6cm). Because of the thickness of the monitor, the displays were elevated approximately 6 cm above the table. The monitor was centered in front of the participant, 6 cm from the front edge of the table. The size of the disks in the displays matched to those in Experiment 1. The displays were presented through Qualtrics (Figure 7). A black three-ring binder was used to block participant’s view of the display between trials.

Design. The design was identical to that of Experiment 1 except noted here. Each participant was assigned, without replacement, to one of the two action task conditions (Touch vs. Grasp), one of two side conditions (the small-disk array on left vs. right), and one of two choice condition (picking smaller disk vs. larger).

In the action response block, participants assigned to the touch task condition indicated their selection of target disks by touching the displayed disks shown on the monitor. Participants in the grasp task condition pantomimed picking up the disk of their selection. They were prohibited to use verbal indication.

Procedure. Procedure was identical to that described in Experiment 1, except that the secondary action tasks were changed.

Data Analysis. Data analysis was identical to those described in Experiment 1. For each set of 12 trials conducted under the same background and task condition, I identified the largest comparison disk selected at least once as being smaller than the standard 28 mm disk. The size of disk was recorded as the criterion value. Effects of condition and response task were assessed using mixed model repeated measures ANOVA. In most cases the sphericity assumption was violated. For all reported results, the Greenhouse-Geisser correction has been applied.
Results and Discussion

Participants’ selection of disk was a function of comparison size in all backgrounds and tasks (Figure 8, Table 2). In general the Ebbinghaus illusion produced a significant increase in criterion value, $F(1, 38) = 275.8, p < 0.0005$. There was no significant interaction between the effect of backgrounds (plain vs. illusion) and response tasks (touch vs. grasp), $p > 0.5$.

A significant interaction between task performance (verbal response vs. action response) and secondary task condition (touch vs. grasp) was identified, $F(1, 38) = 11.7, p = 0.002$. This interaction is difficult to interpret. Because the effect does not interact with the illusion background, however, it is not directly relevant to the hypothesis addressed by this study.

Experiment 2 did not exhibit the reduction of illusion magnitudes observed in Experiment 1 and other previous studies. When the illusion displays and targets disks are shown in 2D format, the participants experience similar magnitudes of illusion in both verbal response tasks and visually-directed action tasks. When participants pantomimed picking up target disks, they experienced no haptic feedback of the true size of the disks, suggesting the haptic feedback is critical in the reduction of illusion magnitudes in visually-directed action tasks.

What’s more, as Vishton et al. (2007) suggested, touching physical target disks reduces the magnitudes of illusion. The touching condition in the current experiment also provided similar size information and sensation experience of the target disks, but the reduction in illusion magnitudes was not present. The difference between the condition in Vishton et al. and the current experiment is that the disks in this experiment were shown on monitors. The lack of 3D size information of target disks may also contribute to the result.
In general, the haptic feedback generated by 3D target disks, and the size information carried by physical objects, seem to play an important role in human size perception in visually-directed action tasks. This information seems to be utilized by the visual processing mode that controls visually-directed action tasks when perceiving object size.

The current experiments examined whether haptic feedback and 3D physical size information contribute to the different performance in different size perception tasks. The next experiment aims to examine this hypothesis by studying the effect of body postures, more specifically hand positions.

**Experiment 3: Effect of Body Posture on Ebbinghaus Illusion Perception**

Previous experiments explored how different visual informations and haptic feedback are utilized by the two visual processing modes. This experiment focuses on another distinct aspect of the two types of tasks: the position of hands. In any visually-directed action tasks, participants inevitably place their hands near the illusion display to execute the touching or grasping action. Such a hand position does not occur in verbal response tasks.

The embodied cognition theory suggests that human cognitive processes, like object recognition, attention, and memory processes, are not purely cognitive. Rather, these processes are tightly linked to control of the body parts, and thus are embodied. For example, memories are suggested to be encoded as their corresponding sensorimotor contexts, body postures, and movements. Recent research has found that, for certain memories, congruent motions and body postures during the memory formation stage and the retrieval stage can facilitate the recall of those memories (Dijkstra, Kaschak, & Zwaan, 2005; Mathôt, Grainger, & Strijkers, 2017; Oakes & Onyper, 2017). For instance, moving up marbles with hands expedites the retrieval of memories with positive valence, as positive memories also convey the sense of going up (Casasanto & Dijkstra, 2010). What’s more, lying down on a recliner
facilitates the recall of personal experience of visiting a dentist, as the body postures in the two conditions are congruent (Dijkstra et al., 2007).

While memory encoding and retrieval are embodied into body postures, object perception is also shown to be related to body postures. Many previous studies have suggested how hand position can alter object perception and recognition, mainly through different allocation of attention (Abrams, Davoli, Du, Knapp III, & Paull, 2008; Gozli, West, & Pratt, 2012).

Abrams et al. (2008) proposed three visual search experiments to explore whether the proximity of hands changes human visual perception. In their first experiment, participants were instructed to find letter H and letter S on a computer monitor and indicate which letter was present. The letters were surrounded by distractor letters. Participants either held the display (so their hands were closer to the stimulus), or put their hands on their laps (so the stimulus was far from the hands). In the first visual search task, participants used their hands to respond. In the second task, they still used their hands to respond, but could not see their hands. In the third task, they responded with their feet. In all three versions of search tasks, participants’ search rate was slower when their hands were near the stimulus than when the hands were far away. This result indicates that the visual processing of stimuli near the hands was slower. More remarkably, similar results were observed when participants were in the hand-proximal condition while their hands were covered by cardboard. When participants knew their hands were near the display but their view of their hands was blocked, they still tended to process the stimulus more slowly than in the hand-distal conditions.

In their second experiment, these researchers tested whether delayed engagement of attention or delayed disengagement affected the result using an Inhibition of Return (IOR) paradigm. When a distractor is at the same position as a later appearing target, people tend to
respond to the target stimulus slower. Such an inhibitory effect of the distractor cue is called the inhibition of return, and it often occurs when the delay is long.

In their experiment set up, a distractor cue was presented at the peripheral of the display. After a short period of delay (delay time varied between trials), the true target appeared. There were still two hand positions (holding the display vs. on the lap). Participants were instructed to report the position of the target. For 300 ms delay trials, reaction times reflected that the distractor cue location captured attention. Hand postures had no effect on reaction times, suggesting that hand positions did not influence participants’ ability to engage attention at cue positions. For 950 ms delay trials, result showed that distractor cue locations inhibited the return of attention. The magnitude of IOR decreased when participants put their hands near the stimuli. In other words, participants responded to the target faster when their hands were near the display.

In their third experiment, these researchers studied how hand postures affect the deployment of attention through attentional blink. Attentional blink happens when two targets appear sequentially. If the second target occurs within a few hundred millisecond after the first one, people tend to ignore the second target. This phenomenon is called attentional blink and it is presumed to reflect the time needed to disengage from the first stimulus. Researchers tested how hand postures might altered attentional blink. Participants were required to identify two specific targets (a number and a letter) from a stream of letters (excluding the ones in target). If hand postures affect attentional blink, then participants who put their hands near the display would be less accurate in detecting the second target. Even though participants in both hand posture conditions showed inaccuracy in detecting the second target, when hands were near the display, the effect of attentional blink was larger, confirming the researcher’s assumption.
The results of Abrams et al. (2008) show that when participants hold their hands near the stimuli, how they allocate their attention to the target display is changed. Participants processed the stimuli more slowly and were slower to disengage attention in both space and time. These effects generated by hand proximity were also present when participants imagined a specific hand posture (Davoli & Abrams, 2009). The visual searching task was the same as the one in the first experiment of the study mentioned above. Before each trial, participants were instructed to imagine either put their hands on the monitor or put their hands behind their back. Even when the postures were purely imaginary, the reaction time was still larger in the hand-proximal condition than in the hand-distal condition. When participants imagined their hands to be near the display, they searched through the display slower than when they imagined their hands to be far from the display.

Studies reported above showed how body postures affect attention. When participants put their hands near the target stimuli or imagine such a posture, they tend to analyze the display longer. Abrams and colleagues (2008) suggested a possible explanation for the change of attention in hand-proximal conditions. Objects near hands are more likely to be manipulated, therefore extended analyses of the object can facilitate the production of movements. In their experiments, hand position changes the proximity of the targets.

The change of attention allocation reported in the studies mentioned above depended on how close the participants’ hands, and thus the participants themselves, were from the target stimuli. When the change of proximity between the participants and the targets is achieved through tool use, how people perceive size and distance information is also changed. Witt et al. (2005) demonstrated that intention to use tools changes how people perceive distance.

In the case of size illusion, when people utilize tools in action tasks, researchers observed reduced illusion magnitudes (Suh & Abrams, 2018). In the first experiment,
participants estimated the size of a target disk by changing the size of a comparison disk (perception task). The target disk was displayed beyond the reach of the participants’ hands, but it could be reached with a stylus. Participants were either instructed to point at the target with their index finger (hand condition) or tap it with a stylus (tool use condition). The target circle remained visible through the experiment. In all cases participants overestimates the size of the target disk, suggesting the existence of a size illusion. For the same disk, participants estimated it as smaller in the tool use condition than in the hand condition. There was no interaction between size and tool use, suggesting the effect was the same across target sizes.

In their second experiment, these researchers increased the number of target disks with more variation of size. The rest of the procedure remained the same. Overall result showed same decreasing trend in tool use conditions, although it did not reach statistical significance. Tool use mainly affected the perception of small target circles. Researchers then changed the stylus to a laser pointer in experiment 3 while other procedures remained the same. They found the same decreasing in perceived size. For the same target disk, participants perceive it as smaller when they reach to it by a laser pointer (within reach) than when they point at it with their index finger (beyond reach of fingers). All three experiments showed consistent results. Post-experiment interviews suggested that the participants were unaware of the true purpose of the study, thus ruling out the possible confound effect of demand characteristic. The observed phenomenon can be categorized as how tool use affected perception, not how tool use affected task response. The results of these experiments suggest that people’s perception on size, which is also an indicator of distance perception, is scaled by action with tool use. And more importantly, the scaling effect depends on one’s ability to act upon the target object.

In summary, allocation of attention as well as object perception can be influenced by body posture, and more specifically hand position. When participants hold their hands near
the target objects, they tend to analyze the object longer and more thoroughly. What’s more, when the target object is within reach of the participants, either by hand or through tool use, and the participants have the intention to act upon the object, participants’ size perception is also influenced.

Previous studies on how visually-directed action affects illusion magnitudes typically involve participants reaching out to perform the visually-guided action tasks. Such an action naturally creates a hand-proximal condition, and confirms the reachability of the objects. On the contrary, in most verbal response task, there is no specific instruction on where participants should put their hands. Participants often tend to put their hands away from the illusion displays, creating a hand-distal condition. In the hand-proximal condition reported in Abrams et al. (2008), participants showed prolonged analyses of the displays. It is likely that in the visually-directed action tasks in previous studies, participants’ hand positions also directed their attention more closely to the displays and induced prolonged analyses of the target disks. Therefore, the reported reduction of illusion magnitudes could possibly arise from such attention allocation and extended analysis. Experiment 3 aims to address this question. If hand position directs people’s attention to the illusion display and generates extended analysis, then simply putting hands near the displays would cause a decrease in illusion magnitudes.

Additionally, Abrams et al. (2008) suggested that the same allocation of attention and prolonged analysis were observed in invisible-hands conditions. When participants put their hands near the display, but their hands were covered so they became invisible, longer reaction time was still reported. Similar result was also observed by Reed, Grubb, and Steele (2006). In their study, hand presence affected spatial attention prioritization and facilitated the target detection near the hand position. Such a facilitatory effect was also observed when the hand
was not visible, but proprioceptive information was available. In the current study, I tested whether an invisible hand-proximal condition would cause a reduction in illusion magnitudes.

Methods

Participants. Seventy-six College of William & Mary students (50 female, 26 male) participated in the experiment for course credit in PSYC 201/202 or as volunteers. Among these participants, seven were excluded because of experimental errors. A total of sixty-nine participants (45 female, 24 male) were included in the final data analysis.

Displays and Apparatus. Displays and apparatus were identical to those used in Experiment 1 except noted here. There was no chinrest or eyepatch in the current experiment. A black three-ring binder was used to block participant’s view of displays between trials. Two cardboard boxes (brown, 40 cm x 15.2 cm x 11 cm, 2 mm thick) were used in the covered condition. The cardboards were arranged in a way that participants could easily fit their forearms into the space created, but their hands were invisible. Two tapes (43 cm from the edge of the table, 4.7 cm from the edge of the display) were used as marks for hand position and cardboard positions. The edge of the cardboards were lined up with the markers (Figure 9).

Design. I randomly assigned each participant, without replacement, to one of two viewing conditions (hand-visible vs. covered), one of two side conditions (the small-disk array on left vs. right), and one of two choice condition (picking smaller disk vs. larger disk). The viewing condition and orientation of display remained throughout the experiment.

Each participants went through 2 blocks of trials. The first block was the lap (hand-distal) condition and the second block was the table (hand-proximal) condition. In 1 block of trials, each participants viewed 6 sets of disk comparisons (27, 28, 29, 30, 31, 33 mm) in each of 4 background conditions (2 plain vs. 2 illusion). The order of comparison pairs and the order of background was randomized across participants. Participants verbally indicated their
selection of disk for all 48 trials. The covering cardboard was visible to participants in the
hand-covered condition throughout the whole experiment.

Procedure. Procedure was the same as in Experiment 1 except noted here. The first 24
trials were identical for all participants. All participants were instructed to put their hands on
their laps during the first block of trials. After 24 trials, participants in the hand-visible
condition (Figure 9) were instructed to put their hands on the table near the display and line
up their fingertips with the edge of the blue marking tapes. They then continued to make the
same verbal judgement as in the first block. Participants in the hand-covered condition were
instructed to put their hands into the space created by the cardboards so their hands were
invisible to them (Figure 9). Then they continued to verbally indicate their selection of disks.

The entire procedure lasted approximately 20 min.

Data Analysis. Data analysis was identical to that in Experiment 1 except as noted
here. For each posture condition (lap vs. table), participants viewed two sets of plain
background displays and two sets of illusion background displays. There were six size
comparisons in each set. The criterion values of each two sets under the same background
conditions were analyzed separately. Effects of body postures (lap vs. table) were assessed
using mixed model repeated measures ANOVA. In most cases the sphericity assumption was
violated. For all reported results, the Greenhouse-Geisser correction has been applied.

Results and Discussion

Participants’ selection of disk was a function of comparison size in all backgrounds
and posture conditions (Figure 10, Table 3). In general the Ebbinghaus illusion produced a
significant increase in criterion value, $F(1, 67) = 305.514, p < 0.0005$. When the two sets of
backgrounds were examined separately, there is a strong posture by background interaction,
$F(1, 67) = 4.136, p = 0.046$. The magnitude of illusion in when participants put their hands on
their laps was 8.7%, while the magnitude of illusion in when participants put their hands on
the table was 7.6%. This represents a 12.6% drop in illusion magnitude. Although the drop was significant, it was much less than the 40% drop reported in Vishton et al. (2007). It was also less than the 20% drop reported in Experiment 1 of the current study.

Additionally, there is a significant set by posture by viewing condition (hand-visible vs. covered) interaction, $F(1, 67) = 5.039$, $p = 0.028$. This interaction is difficult to interpret. Because it does not involve backgrounds (plain vs. illusion), it is not directly relevant to the main question of this experiment.

The experiment presented here focuses on whether hand position affects illusion perception. Results show that when participant put their hands near the illusion display, they experience the illusion differently comparing with when they put their hands on their laps. The position of hands, especially the distance between hands and the display, seem to be sufficient in causing a reduction of illusion magnitude. The reason behind this effect, as suggested by Abrams et al., is that hand position changes the allocation of attention and allows prolonged analysis of the target display. It is worth noting that such a reduction is relatively small, indicating that hand position is only a partial explanation for the performance difference in previous studies.

General Discussion

Researchers have long noticed the difference in illusion magnitudes in verbal response tasks and visually-directed action tasks. People engaging in visually-directed action tasks seem to be more immune to the illusion than when they verbally report their size perception.

One explanation for such a performance difference is the TVSH theory. Some researchers state that there are two separate visual pathways that independently control the size perception processes in the two types of tasks. An alternative explanation is the TMOP hypothesis. Instead of two separate visual pathways, researchers state that there is only one
visual pathway. When engaging in different tasks, the two visual processing modes weigh the same set of visual information differently, resulting in different magnitudes of illusion. The three experiments presented here offered new perspectives to the debate. Using the Ebbinghaus illusion as the paradigm, each experiment focuses on one specific aspect of the visual information that can be utilized in size perception.

Experiment 1 examined the effect of motion parallax and binocular depth cues, the two most effective visual cues in one’s personal space. When these sources of information were eliminated from the viewing conditions, the reduction in illusion magnitudes reported in previous literatures were not present. This result suggests that motion parallax and binocular depth cues are utilized to a similar extent by the two processing modes.

Experiment 2 focuses on the 3D nature of the target disks in the classical Ebbinghaus illusion paradigm set-up. Since the target disks in previous studies were all physical disks, participants could gain haptic feedback in visually-directed action tasks. What’s more, the physical dimensions of target disks might be sufficient in providing size information. The second experiment addressed these issues. When the target disks were presented digitally on a computer monitor, participants reported similar magnitudes of illusion in the verbal response and visually-directed action tasks. The tactile experience generated by graspable 3D disks as targets seem to be critical in reducing the illusion magnitudes.

The third experiment explored whether different body postures in the two types of tasks contribute to the reduction of illusion magnitudes. In visually-directed action tasks, participants have to reach out to act upon the target disks, shortening the distance between their hands and the stimuli. Hand proximity is shown to change the allocation of attention as well as size perception. Experiment 3 tested whether hand proximity is enough to generate the decrease in illusion magnitudes. Results show significant drop of illusion magnitudes when participants put their hands near the stimuli without acting upon the stimuli. Such a
drop is much smaller than reported in previous experiments, however, suggesting that hand position only partially explains the observed shift of illusion magnitudes in visually-directed action tasks. Other factors that may also contribute to the reduction include the haptic feedback from target disks, as illustrated in Experiment 2. Intention to act upon the targets, as suggested by the experiments in Vishton et al. (2007) is also a candidate for reducing illusion magnitudes. Different viewpoints and body size are also possible factors, which I will discuss later.

The experiments presented here provided evidence that physical target disks and hand-near-stimuli positions are critical in reducing the illusion magnitudes. Motion parallax and binocular depth cues, however, do not seem to have critical impacts on the reduction. All three experiments suggest possible modification and directions for future experiments.

In Experiment 1, motion parallax was removed from the perception process through a chinrest. Although the application of a chinrest effectively stabilized participants’ head, it did not completely fixed their head position. There was still space for participants to move their head, which could generate motion parallax. In other words, motion parallax was significantly reduced, but not completely removed from the perception process. Possible improvements in future experiments could address the issue of completely eliminating motion parallax.

For Experiment 2, I presented digital images of the illusion display to participants. Data from the Grasp condition in Experiment 2 suggest that when the targets were not graspable, there was no reduction in illusion magnitudes. Further comparison between these data and the data from the freeview condition in Experiment 1 can provide more detailed evidence for the importance of 3D targets. Moreover, to further investigate the influence of tactile experience generated by target disks, future experiments could present 3D disks on the
computer monitor. Such a combination of target disks and digital illusion display can provide more concrete evidence on the importance of graspable targets.

In Experiment 3, I used the Ebbinghaus illusion paradigm to test the embodied cognition theory. As the embodied cognition theory suggest, body postures, specifically hand positions, can alter how people allocate their attention in during visual perception. In the case of the Ebbinghaus illusion perception, there is evidence that hand position changes attention allocation and causes reduction in illusion magnitudes. In the experiment by Abrams et al. (2008), the direct consequence of the change of attention allocation was a prolonged analysis of the targets. They observed longer reaction time and higher inhibition-of-return in the hand-proximal conditions. In the experiment presented here, I did not record the reaction time of participants. Future recording and analysis on reaction time is necessary to illustrate the effect of attention allocation, which is possibly caused by altering body postures, on illusion perception.

While attention allocation is likely the reason behind the reduction in illusion magnitudes in different hand positions, another possible consequence of altering body postures is the change of viewpoints. According to Wraga, Creem, and Proffitt (2000), another critical difference between various size perception tasks is that people tend to utilize egocentric viewpoints to perceive object size in visually-directed action tasks. In verbal response tasks, people tend to use allocentric viewpoints. In the hand-proximal condition of the third experiment, the targets became more likely to manipulated, as suggested by Abrams et al. (2008). Therefore, when the targets get closer to the participants, participants might view the targets through egocentric viewpoints, even when they do not intend to interact with the targets. In the hand-distal conditions, the illusion display lied relatively far away from participants. Just as in most verbal response tasks, they might view the display simply as an object in a distance, and thus tend to use allocentric viewpoints.
What’s more, in the setup of Experiment 3, I did not control the viewing distance in the two viewing conditions (lap vs. table). In order to put their hands at the designated position on the table, participants often tended to lean forward. Although they were told to sit in natural positions in both viewing conditions, the lack of control in viewing distance might result in a change in viewpoints.

Besides allocation of attention and viewpoints, body postures can also alter human size perception by providing additional size information. Embodied cognition theory suggests that body size information of oneself is used as a metric in size perception (Warren & Whang, 1987; Wagman & Taylor, 2005). Specifically, people with larger bodies would estimate the same width as smaller compared to people with smaller bodies.

In the study by Stefanucci and Geuss (2009), researchers investigated whether body postures affect the perception of aperture width in extra-personal space. They hypothesized that increasing body size through different postures would make the person perceive the same width information as smaller.

In the first experiment, participants saw two wooden poles with different distances in between. They were asked to imagine walking through poles without rotating their bodies. Then they estimated the width between the poles using measuring tapes. Result shows that for the same width, the estimation of broad-shouldered people is smaller than that of the narrow-shouldered people, confirming their first hypothesis.

In the second experiment, these researchers altered the participants’ body width. Participants expanded their body width by either wearing a rod, or holding a rod, or extending their arms. The control group did not extended their body width. All participants imagined walking through several apertures, and then estimated the width of apertures. On average, participants who extended their arms, either holding or not holding a rod, estimated the width to be smaller than the wearing condition and the control condition.
In the third experiment, participants were asked to alter the location of their hands as a representation of different body width. They performed the same tasks as in experiment 2. For each participant, there were a total of 4 arm positions. On average, when participants extended their arms out of their bodies, they judge the same aperture widths to be smaller. When arms were extended, participants estimated the width to be smaller when arms were at the farthest than when arms were closer but still extended. There was no significant difference when arms were not extended beyond body.

The results of Stefanucci and Geuss (2009) show that people’s perception of aperture width changes when their body width is altered through different postures. If people use body size as a measuring metric to perceive size in distant, they might also use body information to judge size in their personal space. Putting hands near the display provided them a chance to do so. Whether body size affects size perception in illusion backgrounds is yet to be determined.

In conclusion, motion parallax and binocular depth cues do not attribute to the drop of illusion magnitudes observed in visually-directed action tasks. The tactile experience generated by 3D target disks and hand position, however, seem to play critical, though only partial, roles in the reduction of illusion magnitudes. Relating back to the broad topic of different perceptual performance in different perception tasks, the three experiments presented here offered new explanations to the cause of the shift in illusion magnitudes. It seems that the different performance cannot simply be ascribed to how the responses are carried out. Rather, variation in external and internal factors (like format of the stimuli or body positions), as results of the different requirements of the tasks, may be critical in the observed performance difference.
References


In Perception of space and motion (pp. 69-117). Academic Press.


Figure 1. Ebbinghaus Visual Illusion
Display of the Ebbinghaus visual illusion. Maximum Grip Aperture is measured through the sensors attached to each participant’s thumb and index finger. Figure reproduced from Vishton et al. (2007) with permission.
Figure 2. Experiment Setup of Loomis et al. (1992)
a. The observer viewed a horizontal target interval, and matched their response (vertical distance, on the sagittal plane) to the target interval. Figure represents the aerial viewpoint. Figures adapted from Loomis et al. (1992). b. Essentially the observer is comparing two depth information, one vertical and one horizontal.
Figure 3. Moving Frame Illusion

a. Participants fixate their eyes on the central dot. The black rectangle represents the background frame.

b. The dashed rectangle represents the original position of the frame. When the frame (solid line rectangle) moves to the left (motion represented by the solid line arrow), participants often report dot moving to the right (represented by the dashed arrow), but the actual location of the dot remain fixed. Figures adapted from Bridgeman et al. (1981).
Figure 4. Illustration of Binocular Disparity

F is the fixation point and f are the foveas. N is the nodal point of the eyes. The black dot will project to different positions on the left and right retina. As illustrated in the graph, the line of projection and lines of sights form different angles, $\alpha$ and $\beta$. Thus the relative retinal locations of the same spot are different. Figure adapted from Howard and Rogers, 1995.
Figure 5.
The illusion display and chinrest used in Experiment 1.
PERCEPTION OF THE EBBINGHAUS ILLUSION

Eye Patch, Chin Rest

No Patch, Chin Rest

Free View
Figure 6. Results of Experiment 1
a-c. Frequency of choosing each disk size as smaller in all tasks (verbal vs. grasp), backgrounds (plain vs. illusion) of all three viewing conditions (Eyepatch chinrest, no eyepatch chinrest, and freeview).
d. Criterion values in different tasks, backgrounds for all three viewing conditions.
Figure 7
The computer monitor in Experiment 2. Both the plain and the illusion backgrounds were displayed on the monitor.
Figure 8. Results of Experiment 2
a-b. Frequency of choosing each disk size as smaller in all response tasks (verbal vs. action), backgrounds (plain vs. illusion) of the two task conditions (touch vs. grasp).
c. Criterion values in different response tasks, backgrounds for the two task conditions.
Figure 9. Two Viewing Conditions of Experiment 3
a. Hand-visible condition
b. Hand-covered condition
PERCEPTION OF THE EBBINGHAUS ILLUSION

Visible Arms

Covered Arms
Figure 10. Results of Experiment 3
a-b. Frequency of choosing each disk size as smaller in two posture conditions (lap vs. table), backgrounds (plain vs. illusion) of the two viewing conditions (hand-visible vs. hand-covered).

C. Criterion values in different posture conditions, backgrounds for the two viewing conditions. Values for each block are plotted separately.
### Criterion Values

<table>
<thead>
<tr>
<th>Viewing Conditions</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Verbal-Plain</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eyepatch+ chinrest</td>
<td>28.37</td>
<td>.891</td>
<td>62</td>
</tr>
<tr>
<td>No eyepatch + chinrest</td>
<td>28.30</td>
<td>.967</td>
<td>73</td>
</tr>
<tr>
<td>Freeview</td>
<td>28.15</td>
<td>.638</td>
<td>59</td>
</tr>
<tr>
<td>Total</td>
<td>28.28</td>
<td>.855</td>
<td>194</td>
</tr>
<tr>
<td><strong>Verbal-Illusion</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eyepatch+ chinrest</td>
<td>31.21</td>
<td>1.243</td>
<td>62</td>
</tr>
<tr>
<td>No eyepatch + chinrest</td>
<td>31.14</td>
<td>1.326</td>
<td>73</td>
</tr>
<tr>
<td>Freeview</td>
<td>31.17</td>
<td>1.275</td>
<td>59</td>
</tr>
<tr>
<td>Total</td>
<td>31.17</td>
<td>1.278</td>
<td>194</td>
</tr>
<tr>
<td><strong>Grasp-Plain</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eyepatch+ chinrest</td>
<td>28.27</td>
<td>.813</td>
<td>62</td>
</tr>
<tr>
<td>No eyepatch + chinrest</td>
<td>28.30</td>
<td>.660</td>
<td>73</td>
</tr>
<tr>
<td>Freeview</td>
<td>28.34</td>
<td>.822</td>
<td>59</td>
</tr>
<tr>
<td>Total</td>
<td>28.30</td>
<td>.759</td>
<td>194</td>
</tr>
<tr>
<td><strong>Grasp-Illusion</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eyepatch+ chinrest</td>
<td>30.68</td>
<td>1.277</td>
<td>62</td>
</tr>
<tr>
<td>No eyepatch + chinrest</td>
<td>30.67</td>
<td>1.313</td>
<td>73</td>
</tr>
<tr>
<td>Freeview</td>
<td>30.63</td>
<td>1.128</td>
<td>59</td>
</tr>
<tr>
<td>Total</td>
<td>30.66</td>
<td>1.242</td>
<td>194</td>
</tr>
</tbody>
</table>

Table 1.
Criterion values for all backgrounds, tasks, and viewing conditions in Experiment 1.
Table 2
Criterion values for all backgrounds, tasks, and response type in Experiment 2

<table>
<thead>
<tr>
<th>Task Condition</th>
<th>Criterion</th>
<th>Std. Deviation</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Value</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verbal-Plain</td>
<td>Touch</td>
<td>28.58</td>
<td>1.017</td>
</tr>
<tr>
<td></td>
<td>Grasp</td>
<td>28.48</td>
<td>.873</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>28.53</td>
<td>.933</td>
</tr>
<tr>
<td>Verbal-Illusion</td>
<td>Touch</td>
<td>31.632</td>
<td>1.3000</td>
</tr>
<tr>
<td></td>
<td>Grasp</td>
<td>31.762</td>
<td>1.1360</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>31.700</td>
<td>1.2026</td>
</tr>
<tr>
<td>Action-Plain</td>
<td>Touch</td>
<td>28.95</td>
<td>.970</td>
</tr>
<tr>
<td></td>
<td>Grasp</td>
<td>28.14</td>
<td>.655</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>28.53</td>
<td>.905</td>
</tr>
<tr>
<td>Action-Illusion</td>
<td>Touch</td>
<td>31.95</td>
<td>1.311</td>
</tr>
<tr>
<td></td>
<td>Grasp</td>
<td>31.14</td>
<td>1.014</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>31.53</td>
<td>1.219</td>
</tr>
</tbody>
</table>
### Criterion Values by Blocks

<table>
<thead>
<tr>
<th>Viewing Conditions</th>
<th>Mean</th>
<th>Std. Deviation</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Set 1</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lap-Plain</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visible</td>
<td>27.900</td>
<td>.8026</td>
<td>35</td>
</tr>
<tr>
<td>Covered</td>
<td>28.029</td>
<td>.9688</td>
<td>34</td>
</tr>
<tr>
<td>Total</td>
<td>27.964</td>
<td>.8842</td>
<td>69</td>
</tr>
<tr>
<td>Lap-Illusion</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visible</td>
<td>30.229</td>
<td>1.5163</td>
<td>35</td>
</tr>
<tr>
<td>Covered</td>
<td>30.426</td>
<td>1.6336</td>
<td>34</td>
</tr>
<tr>
<td>Total</td>
<td>30.326</td>
<td>1.5667</td>
<td>69</td>
</tr>
<tr>
<td>Table-Plain</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visible</td>
<td>27.94</td>
<td>.725</td>
<td>35</td>
</tr>
<tr>
<td>Covered</td>
<td>27.94</td>
<td>.649</td>
<td>34</td>
</tr>
<tr>
<td>Total</td>
<td>27.94</td>
<td>.684</td>
<td>69</td>
</tr>
<tr>
<td>Table-Illusion</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visible</td>
<td>30.114</td>
<td>1.1317</td>
<td>35</td>
</tr>
<tr>
<td>Covered</td>
<td>30.015</td>
<td>1.1579</td>
<td>34</td>
</tr>
<tr>
<td>Total</td>
<td>30.065</td>
<td>1.1373</td>
<td>69</td>
</tr>
<tr>
<td><strong>Set 2</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lap-Plain</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visible</td>
<td>27.971</td>
<td>.7854</td>
<td>35</td>
</tr>
<tr>
<td>Covered</td>
<td>27.721</td>
<td>.6763</td>
<td>34</td>
</tr>
<tr>
<td>Total</td>
<td>27.848</td>
<td>.7392</td>
<td>69</td>
</tr>
<tr>
<td>Lap-Illusion</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visible</td>
<td>30.129</td>
<td>1.2328</td>
<td>35</td>
</tr>
<tr>
<td>Covered</td>
<td>30.206</td>
<td>1.2005</td>
<td>34</td>
</tr>
<tr>
<td>Total</td>
<td>30.167</td>
<td>1.2086</td>
<td>69</td>
</tr>
<tr>
<td>Table-Plain</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visible</td>
<td>28.06</td>
<td>.802</td>
<td>35</td>
</tr>
<tr>
<td>Covered</td>
<td>28.03</td>
<td>.627</td>
<td>34</td>
</tr>
<tr>
<td>Total</td>
<td>28.04</td>
<td>.716</td>
<td>69</td>
</tr>
<tr>
<td>Table-Illusion</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visible</td>
<td>29.900</td>
<td>1.4338</td>
<td>35</td>
</tr>
<tr>
<td>Covered</td>
<td>30.265</td>
<td>1.1094</td>
<td>34</td>
</tr>
<tr>
<td>Total</td>
<td>30.080</td>
<td>1.2880</td>
<td>69</td>
</tr>
</tbody>
</table>

**Table 3**

Criterion values for all viewing conditions, postures, and backgrounds in Experiment 3.