Spatio-temporal integration of an object's surface information in mid-level vision

A DISSERTATION
SUBMITTED TO THE FACULTY OF
UNIVERSITY OF MINNESOTA
BY

Shin Ho Cho

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Professor Daniel J. Kersten and Sheng He

June 2015

Acknowledgements

I would like to express the deepest appreciation to my two academic advisers, Professor Daniel J. Kersten, and Professor Sheng He. Both have the attitude and the substance of a genius: they continually and convincingly conveyed a spirit of adventure in regard to research and scholarship, and an excitement to explore the unknown intellectual world. Without their guidance and persistent help this dissertation would not have been possible.

I would like to thank my committee members Professor Stephen Engel and Professor Bin He, whose comments and encouragement were very helpful in progressing my research and enabling me to finish this dissertation project.

Also, I thank the University of Minnesota for financial support, especially the Doctoral Dissertation Fellowship Program.

Shinho Cho (조신호)

To my parents and wife

Abstract

The human visual system can construct a 3D viewpoint of visual objects based on their 2D contours. This process is presumably an essential part of the visual system that enables interaction with an environment, e.g., grasping an object. Despite its importance, it is not clearly understood to what extent conscious awareness is involved in constructing 3D information from visual input. Here we investigated whether the 3D viewpoint of the object could be extracted and represented by the visual system when observers were not aware of the object's image. To test whether the viewpoint of an invisible cube image could be processed, we measured how much the initial viewpoint of the Necker cube could be biased after adapting to an unambiguous version of Necker cube that rendered. We found that a significant amount of viewpoint adaptation aftereffect occurred even when 1) the adapting cube was invisible due to flash suppression, 2) both the adaptation and test cubes were presented in different sizes and retinal locations, and 3) the adaptation cube was presented to the opposite eye from the target eye. These results suggest that the visual system can construct the representation of a 3D viewpoint in the absence of awareness to visual input. These results are consistent with observations in blindsight patients that appropriate visuo-motor action can be executed with less dependence on the presence of explicit visual awareness.

Our brain resolves the perceptual ambiguity of sensory input. The bistable perception (e.g., Necker cube and binocular rivalry) is a typical example of this perceptual ambiguity. There are many examples of bistable perception and, despite the apparent similarities and differences between them, it remains unanswered whether they are governed by a single neural mechanism for resolving perceptual ambiguity. We measured the switching rates of three bistable perceptions across visual fields (left/right) and eyes (left/right) in both right-handed and left-handed subjects. Results showed that the temporal dynamics of binocular rivalry might be determined by both eye- and hemispheric-specific factors that are dependent on handedness: the Necker cube has a right hemisphere advantage (faster switching) regardless of handedness, and the rotating cylinder has a right eye advantage (faster switching) for right-handed subjects. These results suggest that, for different bistable phenomena, competitions between alternative perceptual interpretations are likely determined by different ambiguity-resolving mechanisms situated along with visual hierarchy.

The human visual system is very good at recognizing and categorizing many different material classes (e.g., wood, stone, metal etc). Glossiness has been known as one diagnostic visual property to judge material class. Many studies found that perceived glossiness was influenced by various intrinsic and extrinsic visual factors, for example, micro-level surface geometry and illumination conditions; nevertheless, it is unknown to what extent viewing time could influence glossiness perception. Here we systematically varied the amount of time for viewing stimuli, and measured how well subjects could discriminate glossy from non-glossy objects. The results showed that perceived glossiness was significantly influenced by viewing time; observers needed at least 300 ms to achieve 75% discrimination accuracy between glossy and non-glossy objects. In further experiments, we also used a rotating object and tested whether rotation speed would be influential to glossiness perception. When the rotation speed was faster, the degree of perceived glossiness became similar between glossy and non-glossy objects. Our findings suggest that glossiness perception is a process to compute the spatial relationship between surface shading information and bright spots, and the efficiency of this computational process is proportional to processing time. A broader implication of the study would be that estimating other material properties (e.g., transparency and translucency) also might be critically influenced by the viewing time.

Keywords: Unconscious processing; 3D information; Viewpoint adaptation aftereffect; Visual field asymmetry; Hemispheric lateralization; Material perception; Glossiness perception; Computational process

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction and Synthesis

Visual perception is a constructive process: the visual system not only registers basic image features such as orientation, spatial frequency, or color, it also interprets the relationship between those features. By integrating such low-level image features, our visual system can produce the representation of more complex object information, such as a 3D structures or surface properties. One intriguing aspect of this process is that it occurs naturally and effortlessly, sometimes even without explicit awareness to sensory input. How can the visual system achieve this? Putatively the idea of "inverse optics" claims that the visual system estimates individual physical elements in the world and reconstructs the scene in the brain from a retinal image. If this proposition is true, then our percept should always be coherent and stable. However, we often experience perceptual ambiguity. The retinal image can be fragmentary, conflicting or even ambiguous. The Necker cube is a well-known example of these ambiguous figures. The viewpoint of the Necker cube is ambiguous because the cube does not provide a 3D depth cue on the image.

One main reason for perceptual ambiguity arising is that the visual system fails to discount individual physical sources integrating them into a single coherent percept. This is reasonable because the retinal image is the outcome of interactions among many physical factors, and discounting individual sources is known to be computationally intractable. For example, the size of an object and the distance between the object and observer are conflated in the retinal image. There are an infinite number of pairings of object size and distance that could give rise to the exact same retinal image. In order to estimate the size of a visual object, the visual system must be able to measure the distance between an observer and the object, which is not always possible.

When visual input becomes ambiguous, our percept often alternates spontaneously every few seconds between two ('bistable') or more ('multistable') interpretations of the sensory input. For example, the viewpoint of the Necker cube is ambiguous, spontaneously fluctuating between two kinds of viewpoints. Alternating percepts can be explained as the process in the visual system to explore the most probable among many possible interpretations, and construct a single stable percept.

Converging evidences have been showing that the visual system processes sensory input putatively through the constructive process. Nevertheless, how this process operates and how attention or conscious awareness influences its operation remains to be answered. One can ask whether the visual system can extract low-level image features (e.g., orientation and contrast). If this is possible, then is it also possible to integrate those features into more complex object features, such as an object's structure, without explicit awareness? In addition, how much time does the visual system need to accomplish this process? Surface properties ("What an object is made of?") are crucial information in recognizing an object. But what if the retinal image changes very rapidly and the visual system cannot extract image information for every single frame? What, then, would our perceptual experience be? We address these questions in this paper and provided some behavioral results.

In this paper we discuss the influence of awareness and viewing time on the processing of sensory information and on reconstructing a scene representation. The paper consists of three relevant studies: in chapter 2 we argue that constructing the representation of an object's viewpoint can occur in the absence of awareness. In chapter 3 we discuss the influence of hemispheric lateralization and the observer's handedness on resolving perceptual ambiguities of bistable figures, and provide an insight on where the underlying mechanism is located in our visual system. Finally, in chapter 4 we look at how the observer's viewing time could influence the processing of an object's surface information; consequently, the perceived glossiness alters depending on viewing time. In chapter 5 we summarize these three studies and consider their further implications to human visual perception in general.

Chapter 2: Processing the 3D structure of an object in the absence of awareness

The human visual system can construct a 3D viewpoint of visual objects based on their 2D contours. This process is presumably an essential part of the visual system that enables interaction with an environment, for example, grasping an object. Despite its importance, we are barely aware of it when 3D information is constructing from visual input; it seems to naturally and effortlessly occur. This raise a question, to what extent conscious awareness is involved in constructing 3D information from visual input.

Here we investigated how the visual system could extract basic-level information of an object (e.g., contours and edges) and integrate that into more complex object properties (e.g. viewpoint) without explicit awareness of visual input. We used the method of interocular suppression (also called continuous flash suppression) to effectively remove awareness to visual input. To test whether the viewpoint of an invisible cube image could be processed, we measured how much the initial viewpoint of an ambiguous Necker cube could be biased after adapting to an unambiguous version of Necker cube. When the adaptation cube was visible, not surprisingly, the adaptors induced a clear viewpoint aftereffect, in that observers were about 20% more likely to perceive the test cube to be the opposite viewpoint of the adaptation cube. Interestingly, a viewpoint adaptation aftereffect was also obtained even when 1) the adapting cube was invisible due to flash suppression, 2) both the adaptation and test cubes were presented in different sizes and different retinal location, and 3) the adaptation cube was presented to the opposite eye from the target eye, indicating that the viewpoint aftereffect could be transferred between the eyes.

Results from this study clearly suggest that the visual system can construct the representation of a 3D viewpoint in the absence of awareness to visual input. These results are consistent with observations in blindsight patients that appropriate visuo-motor action could be executed independent of explicit visual awareness.

2.1 Introduction

The visual perception is a constructive process: the visual system registers basic image features, and also interprets their spatial relationship, constructing a three-dimensional representation of the visual object (Goebel et al., 1998). Although the retinal image is often fragmentary and ambiguous, the 'constructive process' allows our perception to remain unified and coherent. Interestingly, we are usually unaware of its functioning. For example, when we see a face, we often experience the face instantaneously and holistically, without much attention to the shape of individual facial features such as the nose or ears. However, there has been a lack of attention to the relationship between this constructive process and conscious awareness. Can the visual system extract and integrate basic information into more complex object properties even with the complete absence of awareness? Which functions are dependent on the presence of conscious awareness? To answer these questions would provide us a better understanding of the role of awareness plays in visual information processing.

Many studies have proposed effective ways of manipulating observers' awareness to visual input. For example, Continuous Flash Suppression (CFS) made it possible to render visual stimuli in the absence of an observer's awareness (Tsuchiya & Koch, 2005; Fang & He, 2005). Since, compared to other methods (e.g. backward masking, rapid serial visual presentation, and so on), it has the advantages of a longer presentation time and less restriction to stimulus parameters such as size , it has been getting popular as a method that allows researchers to dissociate awareness from physical retinal stimulation. Up to dates, studies using CFS have shown very interesting results. For example, researchers have shown that visual information such as orientation (Pearson & Clifford, 2005), color (Tsuchiya & Koch, 2005), luminance (Harris et al., 2001), contrast (Lehmkuhle & Fox, 1975; Wade & Wenderoth, 1978), and spatial frequency (Tsuchiya & Koch, 2005) can be processed by the visual system even in the absence of awareness to visual input. It has still been a matter of debate, however, whether such low-level information can be integrated and conveyed further along the visual hierarchy, reaching the stage of high level semantic processing (e.g., recognition and categorization). Moradi et al. (2005) showed that facial identity did not generate an aftereffect when the face was invisible to observers due to interocular suppression. There was no indication of semantic words being analyzed on ERP components when the words were invisible (Kang et al., 2011). Invisible pictures could not induce the priming effect in naming the object's category (Cave et al., 1998). All these findings seem to suggest that only certain low-level image features (e.g. orientation) can be unconsciously processed, while semantic meanings of visual input cannot reach the stage of high level semantic processing in the absence of awareness. Other evidence, on the contrary, has shown that many kinds of semantic information can be extracted and processed without awareness. Unlike facial identity, the facial expression eliciting emotional information (e.g. fear) evoked responses in the amygdala (Williams et al., 2004) and significantly greater N170 amplitude in EEG potential (Jiang et al., 2007). Moreover, a fearful face was able to escape from interocular suppression rapidly (Yang et al., 2007). In a similar vein, the subset of an English word could evoke the semantic priming effect even if the test probe word was invisible (Costello et al., 2009). Upright faces emerged from interocular suppression (CFS) sooner than upside down faces did (Jiang et al., 2007). The motor-relevant semantic information (e.g. manipulability) generated a priming effect (Almeida et al., 2008; Almeida et al., 2010) for naming a tool when prime stimuli were invisible (see also, Sakuraba et al., 2010). More recent findings suggest the possibility that some extent of higher order stimulus configuration and semantic meaning can survive in the absence of awareness, and can reach the stage of high-order semantic processing.

Although many studies have shown that high-level information can be processed in the absence of awareness, there are very few studies that examine whether the intermediate level of processing is possible or not. For example, a cube object consists of many edges and surfaces, which are the basic level components that determine its organizational structure. The intermediate level of processing is to combine the low-level information into more complex object properties. Evidently, some studies have shown that complex semantic stimuli (e.g., a face) rendered invisible could evoke the semantic priming effect, indicating that integration of basic visual features can arise even when awareness was removed. However, there is no direct evidence that addresses the possibility of perceptual integration occurring independent of awareness.

To test whether such perceptual integration could occur or not, we measured the viewpoint aftereffect evoked by 2D line-drawing objects (Necker cube) which were rendered invisible. First, the observers adapted to the unambiguous version of the Necker cube while it underwent interocular suppression; in the subsequent test phase, they reported the viewpoint of a bistable Necker cube. The ratio of the viewpoint evoked by the Necker cube was measured through repetitive trials across observers. To render the adapting cube invisible, continuous flash suppression images (CFS) were presented to the other eye while the adapting cube was presenting to the target eye. We predicted that the intermediate process of integrating basic-level object features could possibly occur without awareness, and consequently that the viewpoint could be influenced by the adaptation stimulus.

2.2 Methods

Subjects

The number of Subjects were 20 (Experiment 1), 15 (Experiment 2a), 17 (Experiment 2b), and 15 (Experiment 3). Observers were recruited from undergraduate student pools at the University of Minnesota. All Subjects were naive as to the purpose of the experiment except one, SH (author). Subjects had normal or corrected-to-normal vision and were right-handed. Subjects gave written informed consent in accordance with procedures and protocols approved by the human-subjects review board of the University of Minnesota.

Stimuli and Apparatus

A dynamic noise pattern consisted of randomly generated chromatic texture patterns based on Gaussian distribution (2° x 2°; 75% in RMS contrast; 85 cd/m^2 in mean luminance). Each image was low-pass filtered (<2° cycle per degree in spatial frequency). The refresh rate of the noise pattern was 10 hz throughout the interocular suppression phase. The temporal and spatial frequency parameters were determined individually for each observer prior to the experiment (S/F range 1° to 2°; T/F range 8 to 10 hz) to maximize the time duration of target stimulus suppression.

The test cube was a typical Necker cube that elicited bistable interpretation in two kinds of viewpoints (the vertical and horizontal size, 1.8° x 1.8°; luminance, 40 cd/m^2). The unambiguous version of the Necker cube was used to evoke the aftereffect of viewpoint. To make the cube unambiguous, we removed the vertical lines inside the Necker cube; therefore the cube elicited only one kind of viewpoint (either a look up or a look down) depending on which vertical lines were removed (see Figure 1). The test cube was always presented on the center of a screen, whereas the adaptation cube changed its size (Experiment 2a) and location (Experiment 2b). The fixation red dot always appeared on the center location where the distance from the lower-left and upper-right T-junction is equal. Observers were asked to always fixate on the red dot to prevent the influence of attention on the appearance of each viewpoint.

All visual stimuli were generated by Matlab Psychophysics Toolbox (Brainard, 1997; Pelli, 1997) and a Windows 7 computer. Two 27”-LCD monitors were used (1920 x 1080 at a 140 Hz refresh rate) with a black background at mean luminance (8 cd/m^2). Stimuli were presented on two monitors that were facing each other and fused to each other through two stereoscopic mirrors placed at a 90° angle in relation to each other. Each eye of an observer viewed each monitor from a distance of 100 cm through a stereoscopic mirror on a chinrest in a darkened room.

Procedures

For experiments in the current study, each trial consisted of two phases. In the first phase, observers were adapting to the unambiguous version of Necker cube, and in the second phase, observers were asked to report the initial viewpoint of the Necker cube. Every trial began with a fixation of 1 sec, and then the unambiguous version of the Necker cube was presented to one of two eyes for 2 sec. After the adaptation phase, the blank image was shown for 200 [ms]. There were two conditions of the adaptation phase: one was to present the adaption cube visibly and the other was to present it invisibly. In the visible adaptor condition, observers could see the unambiguous cube, which elicited either 'look-up' or 'look-down' viewpoint. To promote observer's attention to the adaptation cube, observers were asked to report the perceived viewpoint of the adaptation cube. In the invisible adaptor condition, CFS images were presented to the other eye in order to suppress the perception of the unambiguous cube that was presenting to the target eye. During the invisible adaptation period, observers passively viewed the fixation dot without moving their eyes. In order to avoid the abrupt change of interocular suppression, the luminance of the cube was gradually ramped up from 0% to 40% in RMS contrast and remained constant until the end of the adaptation phase. Two experimental conditions pseudo-randomly appeared, and observers performed each condition in separate blocks to avoid task complexity.

In Experiment 1, the typical ambiguous Necker cube was presented to the same eye in the same retinal location as where the adaptation cube was. The observers' tasks were to press one of two buttons to report the initial viewpoint of Necker cube. When the observers viewed the adaptation cube in the invisible condition, they pressed a button (spacebar) to indicate that the dynamic noise did not fully suppress awareness of the adaptation cube. Those trials were excluded from analysis to assure that the adaptation occurred invisibly. The appearance frequency of each type of viewpoint in adaptation cubes was counter-balanced, and their sequences were pseudo-randomized across the experiment. In Experiment 2, the procedure was the same as with Experiment 1, but the size (Experiment 2a) or retinal location (Experiment 2b) of the adaptation cube varied trial by trial. In Experiment 3, the same adaption cubes were used as in Experiment 1, except the adaptation cube and test cube were always presented in different eyes.

2.3 Results and discussion

Experiment 1: The viewpoint aftereffect evoked by the unambiguous cube adaptor

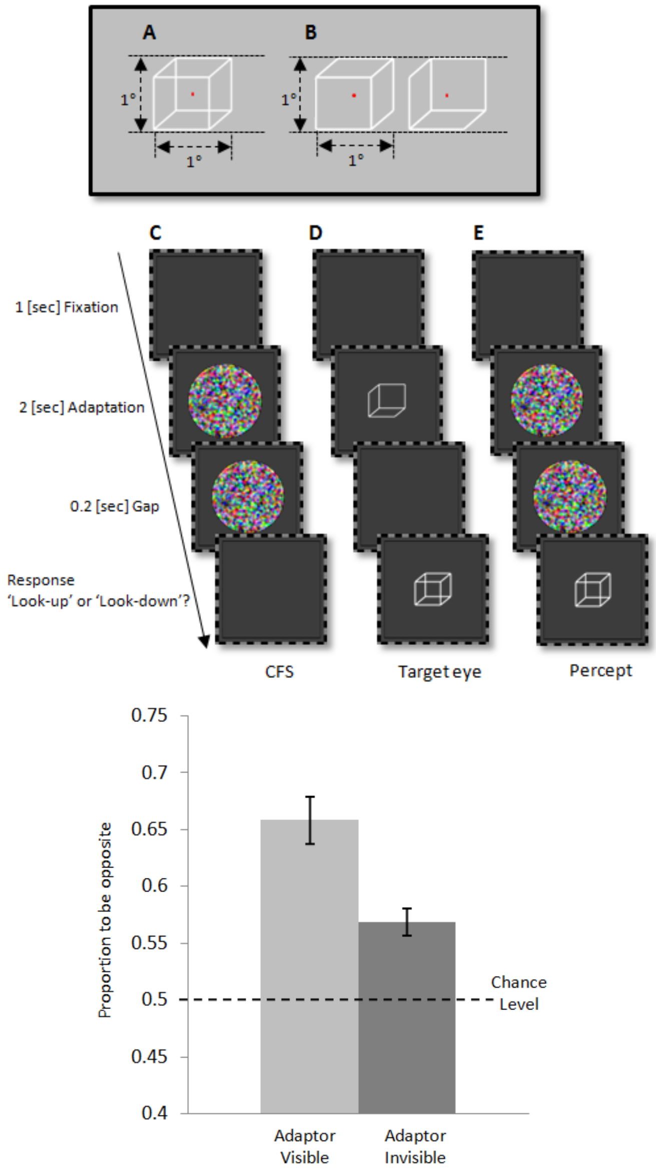

Figure 1. Stimuli, procedure, and results from Experiment 1. The stimuli and procedure used in Experiment 1 (Top/Middle). Bar plot showing proportions of Necker cube to be the opposite viewpoint with the viewpoint of adaptation cube (Bottom). The chance level is 50% based on 2-AFC procedure.

In Experiment 1, we tested whether the visibly and invisibly presented unambiguous version of Necker cube could evoke the viewpoint aftereffect on the Necker cube subsequently presented. We calculated the ratio of the perceived viewpoint of the Necker cube to be opposite the viewpoint of the unambiguous adaptation cube across subjects, and then conducted one sample T-test to figure out whether the probability of the opposite viewpoint was above the chance level (50%) for each visible and invisible adaptor condition. Figure 1 clearly shows that for both the visible and invisible adaptor conditions, observers reported that the initial viewpoint of the test cube (ambiguous Necker cube) was more likely to be the opposite viewpoint that the adaptation cube (unambiguous) showed. In the visible condition, the probability of the opposite view was 15.83% more likely than the chance level (μ=65.83%, σ=.08, one-sample T-test, $t_{14} = 7.65$, $p < .0001$). Interestingly, a similar effect was obtained even when the adaptation cube was suppressed by interocular suppression. Observers were about 6.86% more likely to perceive the opposite viewpoint against the adaptation cube (μ=56.86%), and this probability was also significantly above the chance level (50%); one-sample T-test, $t_{14} = 5.69$, $p < .0001$. These results indicate that the viewpoint of the unambiguous cube evoked an adaptation aftereffect.

The results suggest that perceiving the specific viewpoint of a cube ('look-up' or 'look-down') can induce the viewpoint aftereffect in the subsequent ambiguous Necker cube. This finding is in line with previous research (Carlson, 1953; Harris, 1980; Long et al., 1992). Those studies have shown that the observer, after adapting to an unambiguous version of bistable figures, reported a viewpoint of the test cube to be opposite that of the adaptation cube for a prolonged time, and that the overall alternation rate of the bistable figure was slower than before the adaption phase. Whereas previous studies have shown that the duration of the opposite viewpoint increased due to adaptation of the viewpoint, our results show that the initial percept of the bistable Necker cube more likely biased the opposite viewpoint against the adaptor.

We first report here that the unconscious processing of viewpoint can occur. The results obtained from the invisible adaptor condition clearly show that even without explicit awareness to the viewpoint of the adaptation cube, the interpretation of the test cube can be influenced. Taken together with previous studies showing that processing low-level image features such as orientation and contrast can occur without awareness, our results suggests that the visual can not only register those image features, but also integrate them into more complex image features such as organizational structure. This finding implies that interpreting and constructing 3D organizational structures from visual input can occur independently to visual awareness.

Experiment 2a: The adaptation and test cube were in different sizes

In Experiment 2, we tested whether the viewpoint aftereffect could still occur when the adaptation cube was presented in different sizes (Experiment 2a) and at different locations in the visual field (Experiment 2b). The overall procedures and tasks in both experiments were the same as with the procedure in Experiment 1.

In Experiment 2a, the adaptation cubes were presented in the same location as the Necker cube, but in 30 percent the size was either smaller or larger than that of a test cube. The results showed that both visible and invisible adaptor conditions evoked significant viewpoint aftereffects. Observers more likely viewed the test cube to be the opposite viewpoint (18.92%) than the same viewpoint of the Necker cube in the visible adaptor condition (μ = 68.92%, σ = .11; one sample T-test, $t_9 = 5.25$, $p < .001$). Similar to Experiment 1, the invisible adaptor also evoked the viewpoint aftereffect, more likely to be the opposite viewpoint (4.1%) than the same (μ = 54.1%, σ = .06; $t_{12} = 2.34$, $p < .05$).

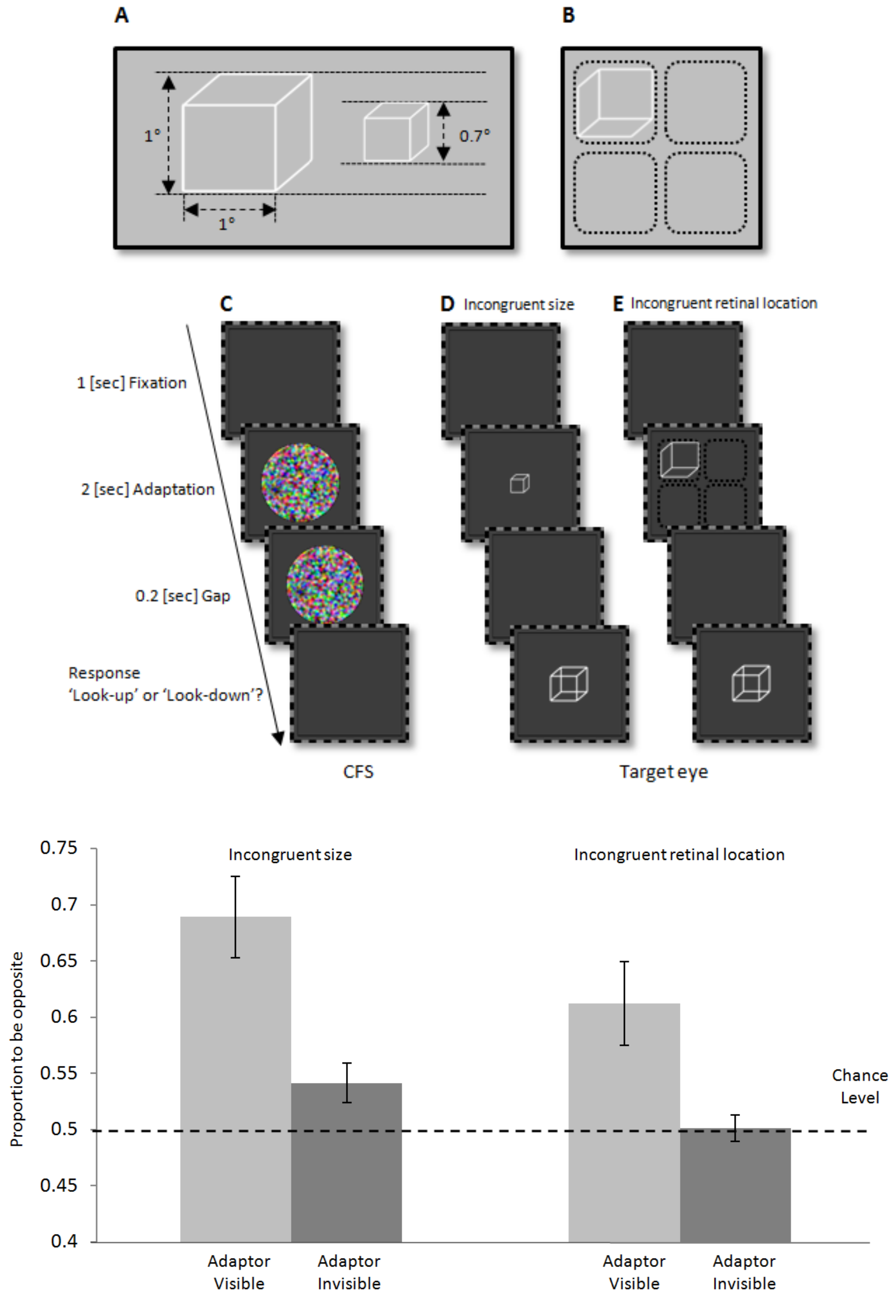

Figure 2. Stimuli, procedure and results from Experiment 2. The stimuli and procedure used in Experiment 2 (Top and middle). Bar plot showing proportions of the Necker cube to be the opposite viewpoint of the viewpoint of the adaptation cube. The two left bars indicate results in the size variation condition in Experiment 2a and the right bars indicate the location variation condition in Experiment 2b (Bottom).

The results indicate that the viewpoint aftereffect is not specific to the size of the adaptation cube. In both visible and invisible adaptor conditions, the measured viewpoint aftereffects were similar to the observed magnitudes in Experiment 1. Because the local image components of an adaptation cube (edges and T-junctions) did not overlap in the visual field with components of the Necker cube, the aftereffect was less likely due to the localized neural adaptation. Rather, the adaptation cube's organizational structure presumably evoked the viewpoint aftereffect. Again, these results clearly support the hypothesis that the visual system can interpret and construct higher order organizational structures in the absence of awareness.

Experiment 2b: The adaptation and test cube were in different locations

In Experiment 2b, we tested whether the viewpoint aftereffect could occur when the adaptation and tests cubes were presented in different locations. We presented the adaptation cubes in one of four locations (left-top, left-bottom, right-top, and right-bottom), but in the same size as the test stimulus (width/height, 1° x 1°). For each trial, the location at which the adaptation cubes were presented was pseudo-randomly decided, and frequency for each location was counter-balanced. There were 60 trials for each location under visible and invisible adaptor conditions; thus, observers performed a total of 480 trials for the experiment. The test cubes were always presented on the center of the monitor through the experiment to prevent possible influence of spatial attention on the perceived viewpoint of the bistable Necker cube.

In the visible adaptor condition, the viewpoint aftereffect was observed (μ = 61.21%, σ = .12; One-sample T-test, $t_{11} = 3.01$, $p < .001$; chance-level = 50%) to be similar to the visible adaptor condition in Experiment 1 and 1-2a. Although the probability of perceiving the opposite viewpoint was somehow reduced, the viewpoint of the adaptation cubes still evoked a significant aftereffect. However, this effect disappeared in the invisible adaptor condition (μ = 50.12%, σ = .04; One-sample T-test, $t_{12} = 0.1$, $p > .1$).

We examined whether the viewpoint aftereffect evoked by the organizational structure of the adaptation cube would be retinotopically location-specific or not. The results obtained in the visible adaptor condition indicate that the effect is not retinotopically specific, since the adaptation cubes presented in the incongruent location to the test cubes still elicited the viewpoint aftereffect, although the overall probability was somehow reduced (61%) compared to the other visible condition in Experiment 1 (65%). Interestingly, when the observers did not perceive the viewpoint of adaptation cubes presented in the incongruent locations, the viewpoint aftereffect disappeared. One possible reason is that observers might have moved their eye fixation to the center of the adaptation cube while it was presenting. In that case, the results of Experiment 2b would be the same consequence showed by the visible condition of Experiment 1. We did not measure the eye movement during the experiment, so this possibility cannot be verified. It is not clear whether the viewpoint aftereffect is specific to the retinotopic location or not. Nonetheless, taken together with Experiment 2a, the results suggest that the viewpoint aftereffect can occur without awareness to the adaptation cubes, and at least that the underlying mechanism might not be sensitive to the size of the adaptor.

Experiment 3: Interocular transfer of the viewpoint aftereffect

In the previous experiment, we examined whether the incongruence of size and location in the visual field between the adaptation and test cubes could change the magnitude of the viewpoint aftereffect. In Experiment 3, we further investigated whether the adaptation cube presented in a different eye from the test eye could induce the viewpoint aftereffect. The stimuli used in this experiment were the same as with Experiment 1, but the adaptation and test cubes were presented in different eyes; for example, the adaptation cube was presented to the left eye and the test cube was presented to the right eye, and vice versa (see Figure 1-3).

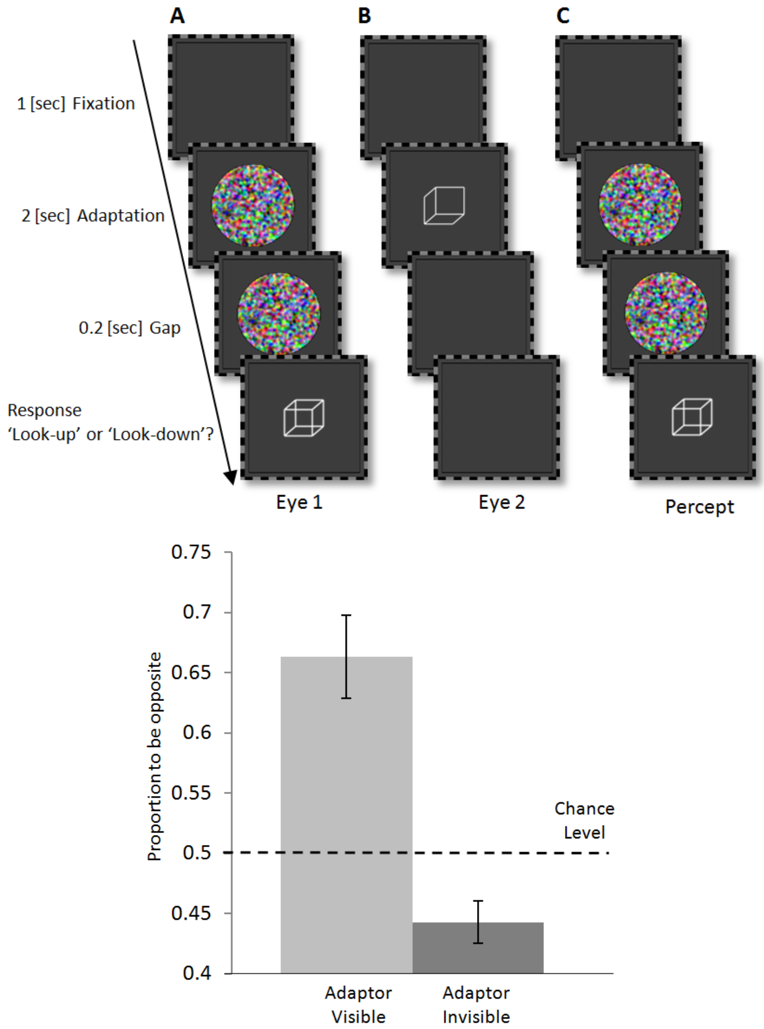

Figure 3. The stimuli and procedure used in Experiment 3 (Top). Bar plot showing proportions of the Necker cube to be the opposite viewpoint of the adaptation cube's viewpoint (Bottom). A proportion of less than 50% means that both the adaptation and test cubes elicit the same viewpoint (Dark-gray bar).

In the visible adaptor condition, we found that the viewpoint aftereffect could possibly be transferred between eyes. The observers were more likely to report the opposite viewpoint of the test cube to the adaptation cube in 16.33%, and the magnitude of bias was similar to the results obtained in Experiment 1 ($\mu = 66.33\%$, $\sigma = .1$; One-sample T-test, $t_9 = 4.72$, $p < .001$; chance-level = 50%). However, when the adaptation cube was perceptually suppressed by CFS, somehow an unexpected result was produced.

The viewpoint of Necker cube was more likely to be perceived as the same viewpoint of the adaptation cube (μ = 44.3%, σ = .07), and this effect was statistically significant (One-sample T-test, t_7 = 2.03, $p < .05$; chance-level = 50%) in Figure 3 (Bottom)

Along with the findings in Experiment 1 and 2, the current experiment indicates that the viewpoint of the visible adaptation cube can produce an aftereffect; this occurs not only when the size and presented location of the adaptation cube is different from the test cube, but also when they are presented in different eyes. The interocular transferred aftereffect implies that binocular neurons would be involved in processing viewpoint of adaptation and test cubes. However, the direction of the aftereffect was reversed compared to Experiment 1. One study conducted by Maruya et al., (2008) showed the reversed aftereffect, which is similar to the effect observed here. They used motion grating adaptors that were presented invisibly, and found the tendency that the motion aftereffect (MAE) which occurred somehow reversed direction in the other eye—however, the effect was not statistically significant.

2.4 General discussion

In this study, we report that the viewpoint elicited by the organizational structure of a 2D line drawing cube image can evoke the adaptation aftereffect. Observers were more likely to report the viewpoint of the Necker cube to be the opposite viewpoint of an unambiguous cube after an adaptation phase for 2 [sec]. Surprisingly, this viewpoint aftereffect was also observed even when the adaptation cube was rendered invisible by presenting interocular suppression images to the opposite eye. These results suggest that the visual system is able to extract and integrate basic image features, such as contour, and construct more complex the 3D viewpoint of an object even without the presence of awareness.

There are studies showing that adaptation can influence the interpretation of bistability on ambiguous figures such as the Necker cube. Prolonged viewing time of bistable images could alter the response profile of retinotopically localized and selective neural channels that are responsible for perceptual switching, resulting in the reversal rate of bistable figure increases or decreases. This phenomenon has been explained by the adaptation of the 'bottom-up' process in the visual system (Spitz & Lipman, 1962; Toppino & Long, 1987; Long et al., 1992; Virsu & Taskinen, 1975). Our results are in line with this idea in that the specific viewpoint elicited by the unambiguous version of the Necker cube increases the probability for the opposite viewpoint to be viewed in the Necker cube. In Experiment 2 and 1-3, the visible adaptation cubes evoked the viewpoint aftereffect even when were presented in a different size, location, and eye from the test cubes. These results clearly suggest that the process mediated by the viewpoint aftereffect is specific to neither retinotopic location nor size. However, these results are inconsistent with some previous findings that showed that adaptation depends on retinal location and size of adaptor. The reversal rate of the Necker cube was negatively accelerated when the cube was viewed for several minutes, however, when the Necker cube moved to a different retinal location the reversal rate returned to baseline (Spitz & Lipman, 1962). Furthermore, the different sizes between adaptation and test stimuli prevented reversal rate 'carry over' from the adaptation phase to the test phase (Toppino & Long, 1987). When the Necker cube was rotating, and thus the local image elements (e.g., contours) were changing their retinal location continuously over time, the reversal rate remained stable rather than being accelerated or decelerated (Howard, 1961). Whereas these studies examined the change of the reversal rates that were measured for a certain amount of time, our studies measured the ratio of the initial viewpoint in the test cube to be the same or different from that of the adaptation cube. This inconsistent result implies that two different underlying mechanisms participate in processing bistable figures: one is to determine the initial interpretation of reversible figures, and the other is to regulate the duration of individual percepts. Specifically the process of determining initial interpretation depends on global and non-location-specific neural channels, while regulating perceptual reversal depends on local and selective neural channels. Although whether or not these two processes are operated by separate neural mechanisms has yet to be answered, the results in our experiment clearly demonstrate that when observers explicitly perceive an organizational structure, organizational information will elicit an aftereffect.

In the invisible adaptor condition in all experiments, the results indicate that, even in the absence of awareness to the adaptor, the visual system extracts and processes the viewpoint of presented stimuli. The organizational structure of the adaptor was not explicitly perceived due to interocular suppression. Nonetheless, the initial viewpoint of the Necker cube in the test phase was biased to the opposite direction of the viewpoint in the adaptation cube. This suggests that the visual system was adapted to the viewpoint elicited by perceptually suppressed structural information of the cube. The viewpoint adaptation by invisible adaptors has not been shown by any studies; thus our study is the first to reveal this effect. Nevertheless, a few studies have found that processing of some intermediate level features could be processed during perceptual suppression. For example, one early study conducted by Oyama & Yamada (1978) demonstrated that the configuration of multiple dots ('horizontally' or 'vertically' elongated) could be estimated even when the stimulus was presented very briefly (50 ms). They found that when the configuration corresponded to the Gestalt principle (e.g. proximity and similarity) the observers could judge its configuration accurately with less viewing time. This result suggests that perceptual grouping can occur very rapidly within 50 ms after the stimulus onset. The unconscious perceptual grouping occurs not only with a target stimulus, but also can evoke a priming effect for the subsequently presented stimulus. In the study of Montoro et al. (2014) the primes consisted of horizontally- or vertically-grouped dot patterns, which was very similar to the stimuli used in Oyama & Yamada's study (1978), and those primes were presented for only 53 [ms]. Interestingly, the results showed that despite observers apparently being unable to perceive the pattern of primes, they showed a faster response time when the global orientation of the test stimulus was congruent with the orientation of prime stimuli. These studies demonstrate that a configuration pattern of visual input (e.g. Gestalt patterns) was registered and processed by the visual system in the complete absence of awareness. We suggest that even before grouping occurs among multiple objects, features of individual object can be extracted and integrated into a more complex single object in the absence of awareness.

Presenting a stimulus for a brief time can effectively remove the observer's subjective awareness to the content of visual input. However, the interocular suppression

(CFS) removes phenomenal awareness, and in addition, the inhibitory process between two eyes' channel disrupts information processing of a stimuli presented in a target eye. Therefore, unlike previous studies, our study indicates that invisible structural information can also survive the binocular inhibitory process. The findings of our study are consistent with one recently conducted study by Wang et al. (2012). By using interocular suppression based on CFS, that study demonstrated that the process of perceptual organization could function even under binocular inhibition. In the experiment, the time for an illusory Kanizsa triangle to emerge from continuous flash suppression (CFS) was measured. The illusory triangle image that was induced by the aligned configuration of inducers ('Pacman') gained awareness faster than the image where inducers were weakly grouped or misaligned. Although it cannot be determined which one was directly responsible for the early emergence - either the grouping of inducers or the presence of illusory triangle - it is a reasonable conclusion that the grouping of inducers occurs before the image gains the observer’s explicit awareness. Taken together with this result, our finding provides evidence that the representation of organizational structure from invisible objects can survive the inhibition of interocular suppression and reach the stage in which the representation of viewpoint in an object is processed.

Interocular transfer of the viewpoint aftereffect in Experiment 3 showed very interesting phenomena because the visible adaptor evoked the same effect originally observed in Experiment 1 and 2, whereas the invisible adaptor evoked the reversed aftereffect that observers were more likely to view the same viewpoint of the Necker cube with the adaptation cube. There have been many studies that show the adaptation aftereffect can transfer between eyes, specifically, that adapting one eye causes threshold elevation in the other eye (Blakemore & Campbell, 1969; Nishida et al., 1994; Falconbridge et al., 2010; Cass et al., 2012). However, there are very few studies that show that the adaptation aftereffect induced by an invisible adaptor can be transferred between the eyes. In the study conducted by Maruya et al. (2008), the motion aftereffect (MAE) inducted by motion grating adaptors was measured. The results indicate that MAE could be transferred between the eyes; however, the magnitude was so weak that it could not register statistical significance. Surprisingly, our results indicate the interocular transfer of the viewpoint aftereffect; however, the aftereffect is reversed compared to other experiments in the presented study. One possible explanation would be that the afterimage of the invisible cube in the adapting eye evokes a facilitation effect for the Necker to be in accordance with the viewpoint of the adaptation cube. Another possibility would be the neural adaptation to one viewpoint in the adapting eye, which facilitates the neural response to the opposing viewpoint in the test eye. Alternatively, the reversed aftereffect may suggest two distinctive underlying mechanisms to process the viewpoint of an object depending on the presence of awareness. However, our data does not tell whether the reversed aftereffect is indeed due to the afterimage in low-level or to other mechanisms in high-level; hence, further experiments are necessary to test these possibilities.

2.5 Conclusion

We tested whether the 3D viewpoint of visible and invisible stimuli could be processed. The 3D organizational structure of adaptation stimuli evoked the viewpoint aftereffect; consequently, the interpretation of the viewpoint elicited by the subsequently presented Necker cube was more likely to be the opposite viewpoint. In addition, this viewpoint aftereffect was not specific to the size, and could be transferred between eyes in both visible and invisible adaptation cube, although the overall magnitude of magnitude produced by an invisible cube was somehow weakened. One major difference of viewpoint aftereffect induced by visible and invisible adaptation cube was that whereas the aftereffect of visible cube was not specific to the retinal location, the aftereffect of invisible cube occurred only when the adaptation and test cube were presented in the same retinal location. Nevertheless, the results clearly showed that the invisible adaptor could elicit significant viewpoint aftereffect.

The results suggest that the visual system can not only extract individual image components, such as contours and edges, but also can combine them to construct a 3D viewpoint even without conscious awareness. Thus we conclude that processing of 3D viewpoint information can occur less dependently of the presence of conscious awareness.

The results of this study provide further implication that presumably other visual processing in the intermediate level of visual hierarchy could occur without awareness. For example, the high-level semantic information, such as the meaning of words or facial expression, could be extracted and processed without awareness. Another implication of the current study would be the unconsciously processed sensory information could influence the observer's behavioral responses related to visual information processing. For example, visually guided viewer-to-object interactions (e.g., reaching and grasping) could arise without observer's explicit awareness.

Chapter 3: The underlying mechanism on resolving image ambiguity and relevance to hemispheric lateralization and observer's handedness: Focused on temporal dynamics.

3.1 Introduction

When visual input produces more than one interpretation, we experience perceptual ambiguity. Perceptual ambiguity often results in our percept alternating spontaneously every few seconds between two ('bistable') or more ('multistable') interpretations of the sensory input. There are many well-known examples: the Necker cube (Boring, 1942), Rubin's face-vase illusion (Rubin, 1958), monocular rivalry and binocular rivalry (Breese, 1909). Various different examples of bistable perception have been discovered. One intriguing aspect of bistable perception is that they have incredibly similar characteristics. For example, in most bistable percepts, the duration of an individual percept cannot be predicted by the duration of the other percept (Levelt, 1967). Although individual duration cannot be predicted, generally, the overall speed of switching percepts (e.g., switching rate per minute) is influenced by some low-level stimulus properties, producing a systematical change of the switching rate. For example, increasing the contrast and luminance (Alexander & Bricker, 1952; Whittle, 1965) or spatial frequency (Blake & Fox, 1974; Wolfe, 1983) of competing images in binocular rivalry yields faster alternation. The switching rate of monocular rivalry is influenced by the orientation and spatial frequency of competing stimuli (Atkinson et al., 1973). Furthermore, competing percepts are usually mutually exclusive to each other, meaning that only one interpretation could be perceived at a time, while others are suppressed from awareness (binocular rivalry) or are less clearly viewed (bistable perception). Perhaps the most striking similarity would be that the individual percepts in different bistable perceptions show a very similar characteristic distribution, which has been known as a gamma (or gamma-like) distribution. This phenomenal similarity naturally poses a question: 'Is there a common single mechanism that drives different bistable perceptions?'

Binocular rivalry has had a long history of study (Blake, 2001), and recently it has become a very popular tool for investigating the neural correlates of consciousness (Tong et al., 2006) because it is able to induce a dissociation between a physical stimulus input and visual awareness to that input. Despite intensive investigations, it is still a matter of debate which specific neural substrate modulates the perceptual switching between two eyes' inputs in binocular rivalry. One recent idea suggests that binocular rivalry involves a distributed cortical network entailing both low-level and high-level processes (Blake & Logothetis, 2002; Freeman et a., 2005), and presumably the interactions between low- and high-level modulates the alternation of percepts. The presence of phenomenological commonality (e.g., the distribution of percept duration) between binocular rivalry and other types of bistable perceptions (e.g., the Necker cube) suggests one common, or at least shared, neural mechanism to modulate perceptual switching in both phenomena (Leopold & Logothetis, 1999; Rubin, 2003; Alais et al., 2000; Kovacs et al., 1996). However, binocular rivalry potentially differs from the Necker cube such that in rivalry there is not only conflict between two interpretations of one sensory input, but also between two different images presented to the two eyes. A Necker cube presented in one eye can still evoke bistable interpretations, suggesting the underlying mechanism presents at higher levels rather than early levels in the visual system. This idea proposes that the key distinction between binocular rivalry and other bistable perceptions like the Necker cube depends on whether binocular inhibition influences or not on perceptual switching.

Here, we investigated whether the switching rates of binocular rivalry and two other kinds of bistable figures (Necker cube, and rotating cylinder) would be different or the same when stimuli were presented in different visual fields and eyes. Further, we also examined the relationship between observers' handedness and the measured switching rates. In particular, it has been found that putatively 'visual field asymmetry’ (VFA). VFA means that when a target stimulus is presented in different visual fields (left or right), their switching rates are differently observed (Chen & He, 2004). The key observation of the current study was to examine whether not only binocular rivalry, but also two kinds of bistable perceptions (Necker cube and rotating cylinder), would show the asymmetrical switching rates in different visual fields. We claims that if both rivalry and the Necker cube have the same asymmetrical switching pattern, then binocular rivalry and bistable perception are driven by one common, higher-level mechanism. Alternatively, if they show different VFA pattern, then different mechanisms would be responsible for each bistable perception. To test this possibility, we presented two different image patterns to the left and right visual fields in corresponding retinal locations in each eye for binocular rivalry. In the following experiments (3 and 4), the Necker cube and rotating cylinder were presented in either the left or right visual field of one eye, with four different retinotopic locations (LE-LVF, LE-RVF, RE-LVF, and RE-RVF).

3.2 Methods

<u>Subjects</u>

Experiment		Right-handed	Left-handed	Total	Average age	Male(Female)
1	Binocular rivalry	8	5	13	24.7	6(7)
2	Interocular suppression	18	8	26	26.5	12(14)
3	Necker cube	18	5	23	25.2	12(11)
4	Rotating cylinder	15	8	23	26.2	18(5)

Table 1. The demographic information of subjects participated in this study. The table shows detailed information of participant demographics in the current study. All observers were recruited from undergraduate student pools at the University of Minnesota. All Subjects except one (author) were naive as to the purpose of the experiment, and all had normal or corrected-to-normal vision. Subjects gave written informed consent in accordance with procedures and protocols approved by the human-subjects review committee of the University of Minnesota.

<u>Stimuli and Apparatus</u>

The diameters of the gratings were 1.8° x 1.8° in visual angle. The circular grating had four cycles and the radial grating had eight cycles. Both gratings were set at full RMS contrast, with the mean luminance at about 15 cd/m^2. In Experiment 1, the stimuli were a red (or green) circular sine-wave grating and a green (or red) radial sine-wave grating. One was presented on the left part of the screen, and the other on the right. In Experiment 2, an achromatic circular grating was used. Depending on conditions, the grating was presented either to the left or right side of the fixation dot, deviated with 1.5°.

A typical Necker cube was used in Experiment 3 (height, 1.8° x width, 2°). The Necker cube was presented on a black background (2 cd/m^2) with a 75% RMS contrast (45 cd/m^2). The cube was located at either the left or right side of fixation (deviation, 1.5°) depending on experimental conditions.

The 2D projection of the cylinder subtended 1.8° (height) x 1.5° (width). The cylinder consisted of 300 small, randomly spaced dots (.08° x .08°). The speed of each dot followed a sine wave function. The dots were white (45 cd/m^2) against a black background (2 cd/m^2).

A dynamic noise pattern consisted of a series of randomly generated chromatic texture patterns (2° x 2°; full contrast in RMS; mean luminance, 45° cd/m^2). The Gaussian random noise patterns were generated and band-pass filtered (spatial frequency < 2 [cpd]). Filtering was performed in the Fourier domain using a 2D Finite Impulse Response filter. The dynamic noise pattern changed every 0.016 [s] (refresh rate, 10 [hz]). The temporal and spatial frequencies were individually calibrated prior to the experiment (s/f range, 3° to 4°; t/f range, 8 [hz] to 10 [hz]) to maximize the efficiency of interocular suppression.

All stimuli were presented through a mirror stereoscope placed at a 90° angle with each other. The mirror stereo scope was mounted on a chin rest. Observers' eyes viewed each monitor from a distance of 100 [cm] through a stereoscope in a darkened room. The monitor was a 27-inch LCD monitor (1920 x 1080 at 140 Hz refresh rate) and all procedures were controlled by Matlab Psychophysics Toolbox (Brainard, 1997; Pelli, 1997). A red fixation dot (.2° x .2°) was always presented at the center of a monitor to promote stable binocular fusion, and Subjects were instructed to fixate on it. A square frame surrounding the stimuli was used to help keep both eyes aligned. The stimuli were presented on a black background with the luminance set to 2 cd/m^2. Each observer was positioned on the chin rest, and the mirrors of the stereoscope were adjusted so that the fixation dot and the frames presented to the two eyes were precisely fused. Each trial began with a central red dot presented to each eye.

Procedure

Experiment 1 (Binocular rivalry) Observers were asked to report their percepts, whether it was radial grating or circular grating, by pressing the left or right arrow button. For each trial observers performed the task for 30 seconds. After taking one-minute practical trials, the observer ran 24 test trials. Each session contained two experimental conditions corresponding to the two visual field positions of stimuli (left and right visual field). For example, when a radial grating was presented to the left visual field of one eye, a circular grating was also presented in the left visual field of the other eye in corresponding location. The occurrence of each type of grating (radial/circular) in each eye was randomized, and the combination of shape and color (e.g. red and green radial grating) was counter-balanced across trials.

Experiment 2 (interocular suppression) At the beginning of each trial, CFS was presented to one of the observer's eyes at full contrast, and the test figure (achromatic circular grating) was presented to the other eye. The contrast of the test figure was ramped up gradually from 0 to 75% within a period of 1s starting from the beginning of the trial, and then remained constant until the observer made a button-press response. A CFS image was presented to both the left and right visual field of one eye, and the test grating was presented to either the left or right visual field of the other eye. To measure the time for a target image to emerge from suppression noise, observers were asked to press the left or right arrow key as soon as possible to indicate in which visual field the test grating appeared. Each participant ran a total of 180 trials (some Subjects had 360 trials). Like Experiment 1, the occurrence rate of the target in the left/right visual field was counter-balanced and randomized in sequence across trials.

Experiment 3 (Reversible Necker Cube) and Experiment 4 (Rotating Cylinder) At the beginning of each trial a 5 [s] fixation dot was presented, and then a typical Necker cube (Experiment 3) and rotating cylinder (Experiment 4) were presented to either the left or right visual field of one eye while no image was presented to the other eye. In Experiment 3, Subjects were asked to report the viewpoint of a Necker cube by pressing the left ('look-up' viewpoint) or right ('look-down' viewpoint) arrow button, and in Experiment 4 the left button (rotating leftward) or right button (rotating in rightward). When the reversal occurred, Subjects were instructed to alter their button press as soon as possible. Each participant performed practice trials for 5 minutes prior to the main test. One trial lasted for 30 [s] and there were 12 or 24 trials total for each participant. The occurrence rate of the target in the left/right visual field and the left/right eye was counter-balanced and randomized in sequence across trials.

3.3 Results and discussion

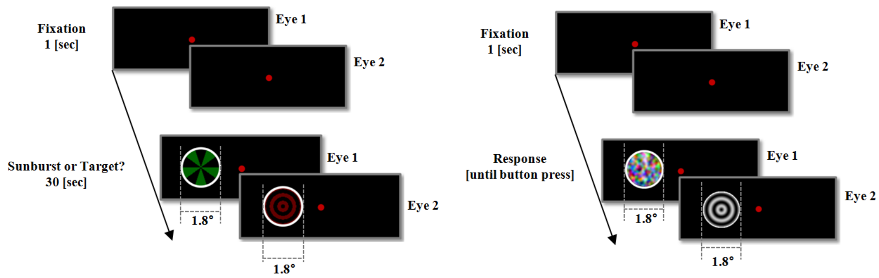

Figure 4. Stimuli and procedure from Experiment 1 and Experiment 2.

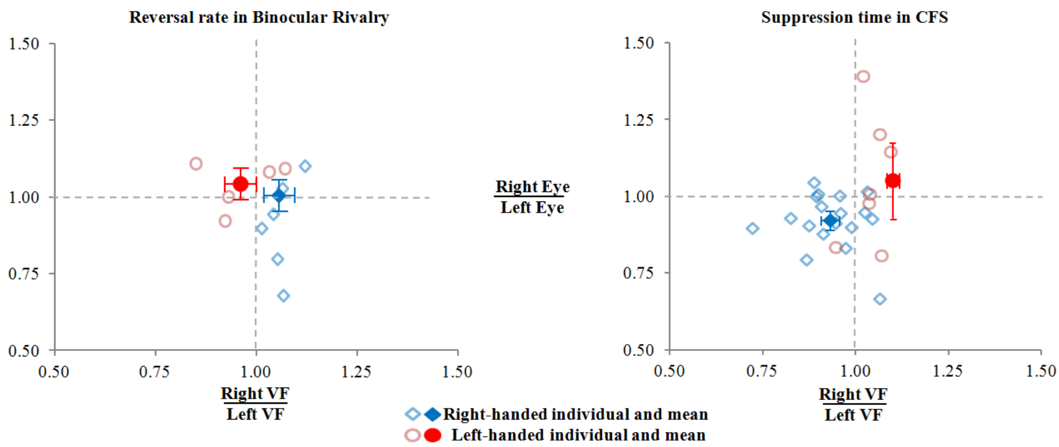

Figure 5. Results of the switching rate ratio from Experiment 1 and 2. The ratio of switching rates of binocular rivalry in the left and right visual field (Left), and the ratio of suppression time in CFS experiment (Right).

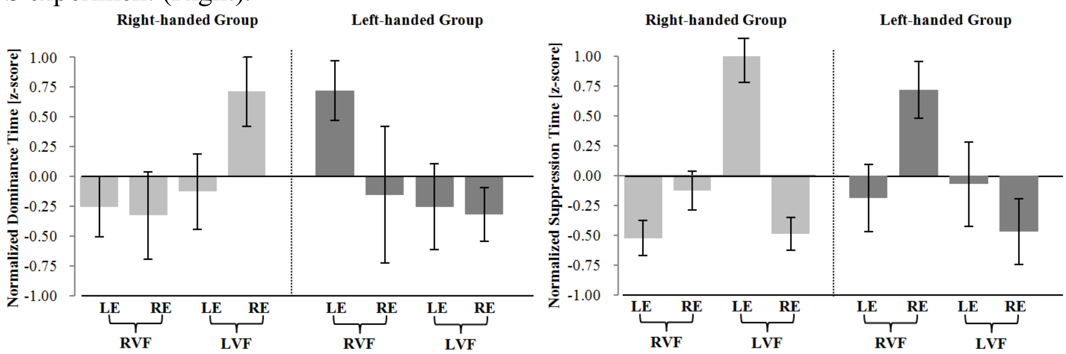

Figure 6. Results of the mean dominance&suppression time from Experiment 1 and 2. Bars indicate the average of normalized dominance time for each grating presented at different retinal locations for different handedness groups (Left). Bars indicate the average of normalized suppression time of a target grating presented at the different retinal locations for different handedness groups (Right).

Experiment 1: Dominance durations in LE-LVF, LE-RVF, RE-RVF, and RE-LVF for binocular rivalry

In Experiment 1 we measured the dominance duration for each competing grating presented in the different visual fields and eyes (LE-LVF, LE-RVF, RE-LVF, and RE-RVF). Figure 2 (Left) shows the ratio between the average dominance duration of each grating that was presented in different visual fields and eyes across different handedness groups. The ratio of individual observers and overall averages between visual fields (x-axis) and between eyes (y-axis) was plotted. Each observer contributed one data point.

For both handedness groups the ratio between gratings presented in each visual field of the left and right eyes was the same, indicating no salient difference between eyes, but we found that the ratio was significantly different between two visual fields for both handedness groups (Figure 4, Left). The right visual field (RVF) showed longer dominance duration than the left visual field (LVF) in the right-handed group, whereas this tendency was reversed in the left-handed group in that the LVF showed longer dominance duration. The average switching ratio between two visual fields LVF and RVF was 1.05 (the right-handed group) and .96 (left-handed group). The ratio for the right-handed group was significantly greater than 1 ($t(6)=4.23$, $p < .01$), however, for the left-handed group the ratio was not significantly less than 1 ($t(4) = 0.976$, $p > .1$). However, the difference between two handedness groups reached statistical significance ($t(6)=2.89$, $p < .05$), suggesting that for each group the pattern of visual field asymmetry is different.

When we plotted each data point separately in the eyes (Figure 6, Left), we found that there was a clear eye dominance effect in LVF for the right-handed group (LVF-RH) and RVF for the left-handed group (RVF-LE), suggesting that an opposite eye dominance effect was contributing to visual field asymmetry in each handedness group. Specifically, for the right-handed group, due to right eye (RE) dominance in LVF on the left eye (LE), the overall switching rate of LVF got slower. For the left-handed group, on the contrary, the dominance of LE in RVF was observed, so that the overall rate of RVF was slower than that of LVF. Moreover, individual data plotted in Figure 6 (Left) shows that most right-handed Subjects had a faster switching rate in RVF, and that most left-handed Subjects showed a faster rate in LVF.

Experiment 2: Suppression durations in LE-LVF, LE-RVF, RE-RVF, and RE-LVF for interocular suppression

In the previous experiment we found that reversal rates in two visual fields were dissimilar, showing visual field asymmetry. To examine which factors resulted in visual field asymmetry, we measured the suppression duration of individual gratings presented in four different locations. We presented a target grating to one eye and dynamic images evoking perceptual suppression to the other eye in corresponding retinal locations for each visual field, one at a time.

Figure 5 (Right) shows the ratio of suppression time between two visual fields (x-axis) and two eyes (y-axis). We found that the gratings presented in LVF, in average, had longer suppression durations compared to the gratings in RVF, in which their ratio (RVF / LVF) was less than 1. Further analysis showed that the target grating presented in LE-LVF took the longest time to emerge from a dynamic noise suppressor (CFS) compared to gratings presented in the other three locations, indicating that the right eye is dominant over the left eye in the left visual field. However, in the right visual field, this tendency was reversed, and the left eye was more dominant than the right eye in RVF. For the left-handed group, interestingly, the suppression durations of each location showed the exact opposite pattern compared to those of the right-handed group. The opposing pattern was in line with the results obtained in Experiment 1. The ratio was larger than 1 (Figure 5, Right), indicating that the suppression in general lasted longer in the RE than the LE, and Figure 6 (Right) shows that the grating in RE-RVF took the longest suppression time in RVF, suggesting that the eye dominance effect of the left-handed group also differed from the effect observed in the right-handed group. To sum up, the data shows that the right eye is dominant over the left eye in the left visual field for right-handed Subjects, and the left eye is dominant over the right eye in the right visual field for left-handed Subjects.

Experiment 3 and 4: Two kinds of bistable figures (Necker cube and rotating cylinder)

In Experiment 3 and 4, we tested two kinds of bistable figures, a Necker cube and rotating cylinder, to determine whether their switching rates would differ between the left and right visual fields. The results were plotted in Figures 8 and 9 in a similar way to Figures 5 and 6.

The most salient feature of the results in the Necker cube experiment is that both the left- and right-handedness groups had a faster switching rate of the Necker cube's viewpoint in LVF. As shown in Figure 8 (Left), the ratio of average switching rates between LVF and RVF was 0.68 (the right-handedness group) and .86 (the left-handedness group), and both ratios were significantly less than 1 (Paired t-test; $t_{17}=7.91$, $p < .0001$ for the right-handed group and $t_3=4.65$, $p < .05$ for the left-handed group). Both handedness groups showed the slowest switching rates on average in the right eye's left visual field (Figure 9, Left). Unlike the results in Experiment 1 and 2, there was no significant difference in switching rates between eyes for both handedness groups.

In the rotating cylinder experiment, however, there was no influence of the visual field on switching rates. Rather, the switching rates of bistable directions on the rotating cylinder showed that the rates were somehow influenced by the eye in which the stimulus presented. However, this eye-specific effect was observed in the right-handed group only. Figure 8 (Right) shows that the cylinder in the right eye elicited faster switching than in the left eye, and the ratio between eyes was 1.08, significantly larger than 1 (Paired t-test; $t_{14}=2.32$, $p < .05$) for the right-handed group. The left-handed group showed a much weaker eye-dependent effect, 1.03, which was not significant ($t_7=0.7$, $p > .1$). However, as noticeable in Figure 9 (Right), the left visual field of the left-handed group showed a large difference of switching ratios between eyes presented.

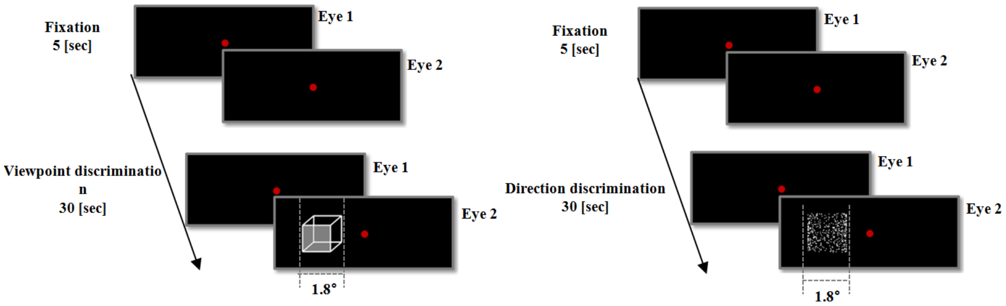

Figure 7. Stimuli and procedure from Experiment 3 and 4.

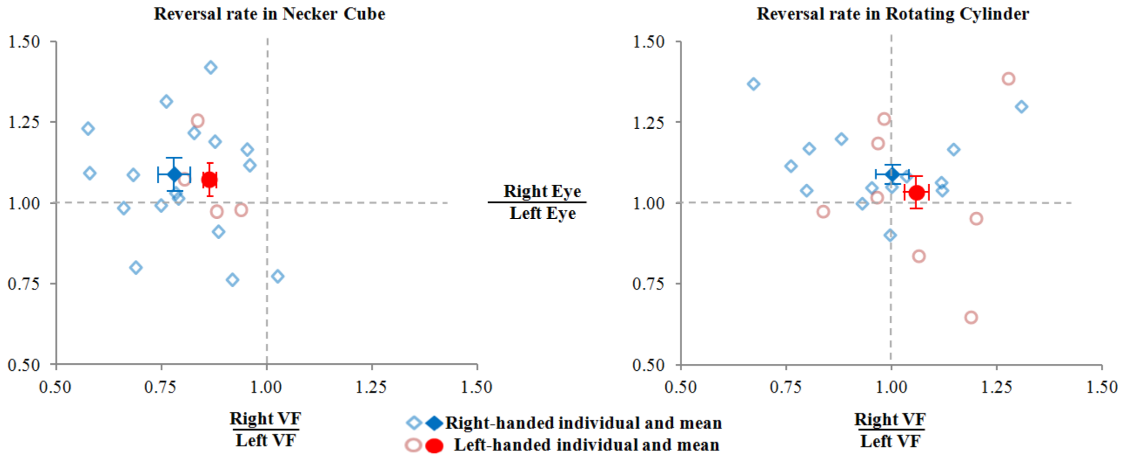

Figure 8. Results of the switching rate ratio from Experiment 3 and 4.The ratio of switching rates of the Necker cube (Left) and rotating cylinder presented in the left and right visual fields (Right).

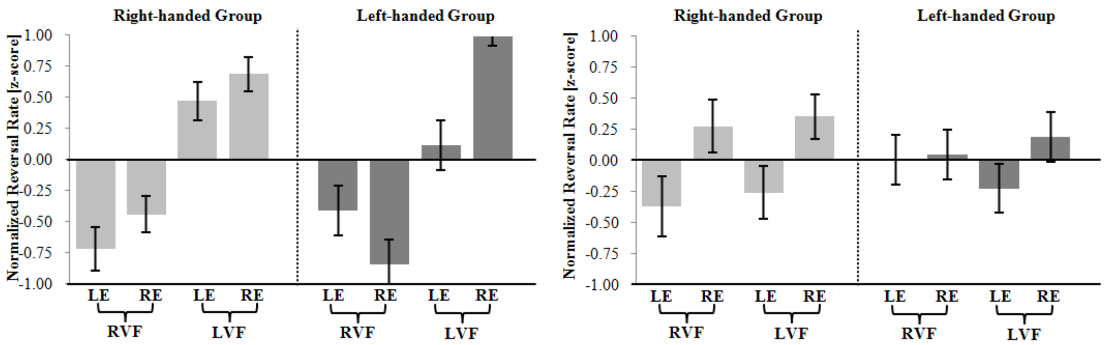

Figure 9. Results of the mean reversal rate from Experiment 3 and 4. Bars indicate the average of normalized switching rates of the Necker cube presented at four locations (LE-RVF, LE-LVF, RE-RVF, and RE-LVF) for each handedness group (Left). The average of the normalized switching rates of the rotating cylinder presented at four locations for each handedness group (Right).

3.4 General discussion

		Right-handed group			Left-handed group		
		R/L Ratio		Which one faster?	R/L Ratio		Which one faster?
		Mean	Std.		Mean	Std.	
Rivalry	VisF	1.055	0.035	RV(**)	0.961	0.089	LV(*)
	Eye	1.002	0.287	RE(n.s.)	1.041	0.078	RE(*)
Necker Cube	VisF	0.781	0.149	LV(**)	0.864	0.058	LV(**)
	Eye	1.088	0.206	RE(*)	1.070	0.132	RE(*)
Cylinder	VisF	1.003	0.220	RV(n.s.)	1.060	0.150	RV(*)
	Eye	1.089	0.148	RE(**)	1.034	0.238	RE(n.s.)

* $p<.05$, ** $p<.005$

Table 2. Summary table of the average switching ratio between visual fields and eyes from in this study.

For both handedness groups, switching rates of rivaling stimuli presented in an individual visual field were not equal in the left and right visual fields. In Experiment 1, two visual fields had significantly different switching rates for both handedness groups; the right visual field (RVF) showed a faster switching for the right-handed group, and the left visual field (LVF) was faster for the left-handed group. The switching of the Necker cube which presented in one eye (Experiment 3) also showed that its switching speed was different across visual fields, but there was no difference between handedness groups; in both groups, LVF was faster in switching rate than RVF.

One possible reason for the different switching rates in different visual fields is that one hemisphere has generally faster neural dynamics than the other hemisphere. Significant left-right hemispheric difference has been demonstrated in many studies of cognitive domains (for reviews, see Hellige 1993; Hugdahl & Davidson 1994) such as language (Rasmussen, 1977; Knecht, 2000; Zhou, 2010), attention (Heilman & Van Den Abell, 1980; Robertson & Ivry, 2000), emotional information processing (Bryden 1982, 1983), spatial information processing (Corballis & Sergent 1989; Roth, 1998; Iachini, 2009), and some neurological disorders (Galaburda et al., 1990; Ross & Monnot, 2008; Toga & Thompson, 2003). In visual perception studies, many studies reported that the right visual field that projects to the left hemisphere has an advantage in processing fast temporal events (Nicholls, 1994; Okubo & Nicholls, 2008). On the contrary, there is also evidence showing that the left visual field that projects to the right hemisphere has an advantage for other cognitive processes. For example, a face perception resulted in greater fusiform activations in the right hemisphere than in the left hemisphere (Kanwisher et al., 1997; Bentin et al., 1996). The response time of simple dot detection tasks was faster when the target dot was presented in LVF (Davidoff, 1977). The LVF had an advantage in spatial phase discrimination tasks (Beradi & Fiorentini, 1991) for grating-type stimulus. All these studies demonstrated that, in general, the right visual field (left hemisphere) is dominant for temporal information processing, and the left visual field (right hemisphere) is dominant for spatial information processing.

In our study, the switching rate of binocular rivalry in RVF was faster than in LVF for the right handedness group, as shown in Table 2. This result seems to be consistent with the idea that the left hemisphere has a superior temporal processing ability. Presumably different switching rates in two visual fields might be mainly due to the influence of hemispheric lateralization. However, a faster switching rate in RVF was only observed for the right-handed group, and on contrary, for the left handedness group LVF showed a faster switching rate, indicating that our results can be partially explained by the idea of hemisphere lateralization.

Alternatively, when we measured the suppression duration of different visual fields and eyes, we realized that the results suggest another possibility to explain the asymmetrical characteristic of switching rates. Only one eye for each handedness group had unusually longer predominance time in their ‘slow’ visual fields (LVF for right-handed, and RVF for the left-handed). Furthermore, prolonged suppression time was observed in LE-LVF for the right-handed group and RE-RVF for the left-handed group. The different dominance of nasal and temporal hemi-retina could account for the difference of switching rates of visual fields. Specifically, for the right-handed group, the right eye’s temporal retina (RE-LVF) with an uncrossed pathway to the right hemisphere could suppress the stimulus for longer duration compared to the opponent eye. For the left-handed group, the left eye’s temporal retina (LE-RVF) had better inhibition in binocular rivalry. The results are well consistent with the idea of nasal/temporal pathway asymmetry. The asymmetrical superiority of Nasal/Temporal retina has been reported in previous studies. The nasal retina is relatively more sensitive to processing the orientation information of stimulus than temporal retina (Paradiso & Carney, 1988). The stimulus received by the nasal hemi-retina, in general, predominated over the stimulus received by temporal hemi-retina for longer time (Fahle, 1987; Fahle & Schmid, 1988). The anatomical investigation of retina showed that the nasal hemi-retina has a higher cone density (Curcio et al., 1987), suggesting the higher cone density might be responsible for the superiority of nasal retina. However, our experiment showed that the temporal retina, 'inferior' in these earlier studies, enjoyed longer dominance time over the opponent eye's nasal reitna in binocular competition. While the current results are not consistent with some previous results, there have been other earlier results that show consistent results. Kaushall (1975) showed that the uncrossed visual pathways (temporal hemi-retina) were dominant over the crossed pathways (nasal retina) in a binocular rivalry task. More recently, Chen & He (2004) found the superiority of temporal retina, which is exactly consistent with the current results. It is not clearly known why the current and previous studies show inconsistent results; one possible explanation might be that the nasal/temporal superiority could be influenced by the position of rivaling stimuli from a central fixation. For example, one earlier study (Fahler, 1987) used the stimuli were presented in a far periphery (30 degree) area, but in our study the stimuli were presented in near fovea area deviated only 1.5 [deg] from the central fixation.

If one common mechanism which drives the perceptual switching underlies both rivalry and the Necker cube, then the pattern of visual field asymmetry in switching rates also must be similar. However, unlike binocular rivalry, there was no salient nasal and temporal pathway difference. Rather, both handedness groups showed faster rates in one visual field: LVF. These results suggest that the modulation mechanism of rivalry might differ from the mechanism of bistable figures. We conclude that binocular rivalry occurs between Nasal/temporal pathways, while Necker cube switching more likely reflects the modulation from a cortical origin in hemisphere.

Our study also shows that handedness influences the asymmetrical switching rates in binocular rivalry. The superiority of temporal retina to suppress the opponent visual input was observed in one eye of each handedness group. However, interestingly, the dominant eye between the two groups was not same; for the right-handed group, RE had superior temporal retina, whereas for the left-handed group, LE had the superior temporal retina. The results showed that functional asymmetry of nasal/temporal pathways might be opposed to each other depending on left or right-handedness. However, the influence of handedness on hemispheric lateralization has been controversial. Some studies have proposed that right-hemisphere lateralization of face processing is independent of handedness (Hamilton & Vermeire 1988), and left-handers may not have a left-hemifield bias for faces (Gilbert & Bakan 1973; Levy et al. 1983; Hoptman & Levy, 1988; Luh et al. 1994). Some studies showed mixed evidence (Borod et al. 1990), suggesting that the influence of observers' handedness on eye or visual field asymmetry is weak. On the contrary, other studies have found that clear, distinctive performance in certain visual fields was observed between different handedness. For example, saccadic movement latencies in right-handers were found to be shorter when shifting to rightward fixation, rather than leftward (Rayner, 1978; Pirozzolo, 1979; Pirozzolo, 1980; Hutton, 1986), and some left-handers showed the opposite tendency, with more rapid leftward saccadic movement. In a similar vein, human neuro-imaging studies have shown the reverse lateralization of hemispheres between different handedness: extrastriate body area (Downing et al., 2001), fusiform body area (Peelen & downing, 2007), human motion area MT and human middle temporal (Zeki et al., 1991; Tootell et al., 1995; Dumoulin et al., 2000), and the two left-handed subjects with apparent left-hemisphere lateralization of FFA (Kanwisher et al. 1997). Despite many findings suggesting that handedness influences hemisphere lateralization, this idea cannot fully account for our data. Both handedness groups showed the superiority effect in their temporal retina in Experiment 1; therefore, the lateralization effect seems to be specific to the eye, rather than brain hemisphere. In Experiment 3, because both handedness groups commonly showed faster switching rates in one hemisphere, the handedness is less likely related with asymmetrical switching rates. We do not rule out the possibility that handedness influences hemispheric lateralization, but we conclude that handedness influences the superiority of nasal and temporal visual pathways for binocular inhibition.

3.5 Conclusion

We investigated the asymmetrical switching rates of visual stimuli presented in different visual fields to figure out whether binocular rivalry and other kinds of bistable phenomena have similar underlying mechanisms or not. We found that binocular rivalry and the Necker cube showed different patterns of asymmetrical switching rates, depending on visual fields and eyes. These results indicated that 1) differing superiority of nasal and temporal visual pathways was more likely involved in the switching of binocular rivalry, and 2) observers' handedness may influence this superiority. Further experiments (Necker cube and rotating cylinder) suggest the bistable perception is more likely influenced by high-level processes (presumably hemispheric level) rather than early-level processes of visual pathways. We conclude that binocular rivalry and two bistable perceptions have different switching mechanisms in visual hierarchy.

Chapter 4: Processing the material property of an object surface within a limited amount of viewing time: Focused on glossiness perception

4.1 Introduction

Studying material perception aims to understand how we can infer the material of an object. A human being can distinguish, without much effort, numerous different materials (plastics, woods, stones, liquids, and organic matter) from their visual qualities. Material perception can be achieved for an incredibly short time, sometimes even without explicit knowledge of an object's identity (Sharan et al., 2008; Sharan et al., 2009, Sharan et al., 2014). Sharan et al., (2014) found that observers could identify material categories in 40 [ms] exposures with 80.2% accuracy. It is useful, sometimes critical, to recognize what objects are made of. For example, one must determine the edibility of a food, or identify a foothold to step. However, it has remained unclear how many kinds of visual properties of material (e.g., glossiness, transparency, roughness, and so on) we perceive from the environment, and what are the most important properties in categorizing materials as a certain class (e.g., wood, stone, metal, and glass).

Glossiness is the sensation that the surface of an object is reflecting light. The typical appearance of gloss is a bright spot that we usually call a 'highlight', or an image reflecting the surrounding environment. Because each material class carries its own characteristic glossiness due to different reflectance and transmission properties (Figure 10), glossiness can be one of the most diagnostic visual properties for judging material class. On one hand, the appearance of gloss has been well described by sophisticated physics models. For example, the Bidirectional Reflectance Distribution Function (BRDF) proposed by Nicodemus (1965) allows us to estimate the proportion of light reflected on the local surface from the overall amount of light arriving on a surface. More recent reflection models have provided more realistic simulations of how glossy appearance looks to the human visual system under varying illumination conditions and with different material properties (Ward, 1992; Oren & Nayar, 1994; Günther et al., 2005).

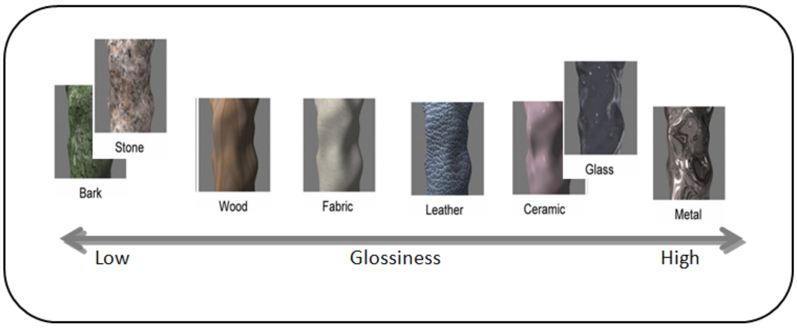

Figure 10. The perceived intensity of glossiness evoked by different kinds of material. Metal looks glossier than material such as stone or bare wood. The demonstration shows that perceived glossiness could play a pivotal role in differentiating numerous material classes.

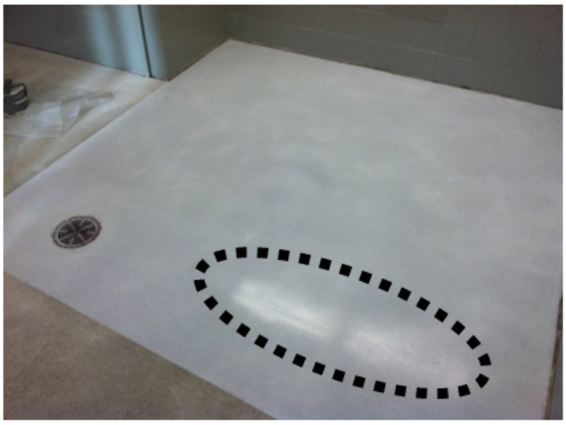

Figure 11. The example of a picture showing an ambiguous brightness. The dotted oval indicates bright spots on a floor. The visual system does not always interpret the bright spot on a surface as highlights.

On the other hand, although those models are able to simulate optical physics and consequently predict how reflection looks to the visual system, it does not tell us how we experience gloss sensation from the consequence of physical interactions. In Figure 2 we can see a bright spot on a floor (black dotted circle); however, the bright spot looks ambiguous because we could not determine whether it is a reflection or a white stain on the floor (e.g., paint or a scratch). This image demonstrates that although brightness must be one important factor in evoking a glossy appearance, the visual system does not always interpret the bright spot as the consequence of light reflection. The physics model can explain the underlying causes of surface gloss; however, the model does not tell us how the glossy sensation can arise from those physical causes.

Many intrinsic and extrinsic properties of an object determine the degree of perceived glossiness, such as the observer's viewing direction (Obein et al., 2004), the shape of the surface (Ho et al., 2008; Nishida & Shinya, 1998; Wijntjes & Pont, 2010), and the illumination field (Doerschner et al., 2010; Fleming et a., 2003; Olkkonen & Brainard, 2010). Perceived gloss increases when the surface has motion properties that change image information over time (Sakano & Ando, 2010; Wendt et al., 2010) and when binocular disparity presents (Wendt et al., 2008). More recently, Motoyoshi et al. (2007) and Nishida & Shinya (1998) proposed an idea that measuring simple image statistics enables one to predict the degree of perceived glossiness. Specular reflections typically generate positively skewed luminance histograms – putatively 'sub-band skewness' – and the visual system computes this statistical property to infer surface gloss. Furthermore, when they modulated the skewness of an image histogram, they found that perceived gloss also systematically varied. However, in contrast, Anderson & Kim (2009) argued that the highly correlated relationship between perceived gloss and histogram skewness is only limited to a restricted set of stimuli under a certain condition of surface geometries, surface reflectance, and illumination fields. They claimed that because the intensity histogram of an image does not reflect the geometrical information of a surface, image histograms cannot distinguish gloss from other possible sources that caused histogram skewness, such as pigmentation, pr surface geometry (Anderson & Kim, 2009; Wijntjes & Pont, 2010).

Indeed the relationship between gloss and surface geometry has been revealed to be crucial to perceived glossiness. The location of the appearance of specular highlights has a strong relation to surface shape (Beck & Prazdny, 1981; Anderson & Kim, 2009, Todd et al., 2004). These results suggest that the visual system may compute the spatial relationship between 3D shape information reconstructed from diffuse shading profiles, and luminance maxima caused by specular reflection. Blake & Bülthoff (1990) showed that the visual system was readily able to infer the proper depth of highlights on a surface curvature, suggesting that the visual system may run a kind of physical simulation to evaluate whether highlights on a surface physically correspond to surface shape. When the location of specular highlights was forcefully displaced to other incongruent locations by using image processing techniques, the perceived gloss diminished as a function of location deviation (Anderson & Kim, 2009). Based on this result, Anderson and his colleagues proposed two factors that strongly influenced the perceived glossiness: orientation and brightness congruence between specular highlights, and surface shading of an image. Perhaps it is not possible to deconstruct all of the physical sources that result in an image, but those results show that the visual system is able to reconstruct surface shape from a surface shading profile and compute the spatial congruence of bright spots with surface geometry in order to estimate the intrinsic reflectance of a surface.

Although many factors that potentially influence perceived glossiness are known, how those factors are processed in temporal domains is unclear. One of the most salient distinctions between the laboratory environment and the real world is that an object in the real world is not static. The object in an image continuously changes its visual appearance (based on size, viewpoint, and location) due to the object's and observer's motion. In this case, the visual system requires a limited amount of time to process image information and to integrate changing information in order to properly estimate material properties. For example, Doerschner & Kersten (2011) showed that specular highlights sliding over a surface elicited a characteristic visual pattern of optic flow, and the visual system was very sensitive to detect these features when judging an object's shininess. However, it is still largely unknown how much time is necessary for the visual system to extract gloss features and how fast the visual system achieves this process. Early studies showed that recognizing object and scene can be achieved very rapidly, even as observers were aware of the presence of an object (Grill-Spector & Kanwisher, 2005). The speed of recognizing natural objects and natural scenes is incredibly fast, even when observers have never seen them before (Thorpe et al., 1996; Oliva & Torralba, 2001; Greene & Oliva, 2009). A typical scene fixation of 275 to 300 ms is sufficient to understand the gist of an image, and observers were able to report the semantic information (e.g., 'a birthday party') of the scene (Intraub, 1981; Tatler et al., 2003). The time required for object recognition depends on the size of an object (Fei-Fei et al., 2007), and different tasks (e.g., localization of an object in a scene) require more exposure time (Evans & Treisman, 2005); nevertheless, the overall time required to perform the task was reasonably fast within 300 [ms]. These studies demonstrate that the visual system can not only extract image features from visual input, but also process those features to produce semantic information in a very short time. So a natural question arises: How long does it take to recognize material from the surface of an object? For example, when a bright spot appears on a surface, how fast can we judge it by specular highlights or white pigment on a surface? How much surface information (e.g. glossiness) can be extracted in the first few hundred milliseconds of seeing an object? Answering these questions can provide a key to understanding whether the underlying mechanism for material perception would be different or the same as other kinds of perception (e.g., object recognition). Specifically, the time course of material perception can suggest whether or not material perception is mediated by the common underlying mechanisms of other kinds of perception.

To investigate how fast the visual system processes a surface gloss, we measured the discrimination accuracy of a glossy surface from a painted matte-like planar surface under varying amounts of a viewing time. To generate a painted matte-like object with a bright spot on it, but without the surface exhibiting gloss, we manipulated its specular highlight layer in a way that was originally proposed by Anderson and his colleagues (2009). To test the influence of stimulus duration, we presented two images (a target glossy surface and a matte surface) in sequence for various amounts of time, and asked subjects to choose which one looked glossier (2-interval forced choice task).

4.2 Methods

Subjects

In Experiment 1, the Subjects were 14 undergraduate students at the University of Minnesota. All Subjects had normal or corrected-to-normal vision and were right-handed (aged 19-32 years, 12 female) and were naive as to the purpose of the experiment.

Subjects provided informed consent in accordance with procedures and protocols approved by the human-subjects review committee of the University of Minnesota. Three subject data sets from individual experiments were excluded from analysis because they could not recognize a gloss on a target surface.

Sixteen undergraduate students at the University of Minnesota participated in Experiment 2. All Subjects had normal or corrected-to-normal vision, were right-handed (aged 19-32 years, 10 females), and were naive as to the purpose of the experiment.

Materials and Stimuli

All visual stimuli were generated using the Matlab Psychophysics Toolbox (Brainard, 1997; Pelli, 1997) and presented on a 27-inch LCD monitor (1920x1080 at 140 Hz refresh rate) connected to a Windows 7 computer. Observers viewed the monitor from a distance of 100cm, with a chin rest. Stimuli were always presented against a black background at mean luminance (8 cd / m^2).

Planar surface

The planar surface looks like an egg carton. The mesh grid was generated by Matlab, and then rendered by 3D computer graphics software (Blender). The surface exhibited complex curvatures based on roughness parameters to determine how many sine waves needed to be synthesized. The surfaces had a square based on the x- and y-plane (width and length respectively) and the z-plane (height component).

Each surface consists of at least fifteen, and up to twenty, randomly oriented sine gratings, which were accumulated along the z-plane (Equation 1). To make a homogeneously rough surface, a high-pass filter was applied after the sine wave grates were synthesized. The height components of the curvature of all stimuli were statistically independent in order to perceive shape differently enough.

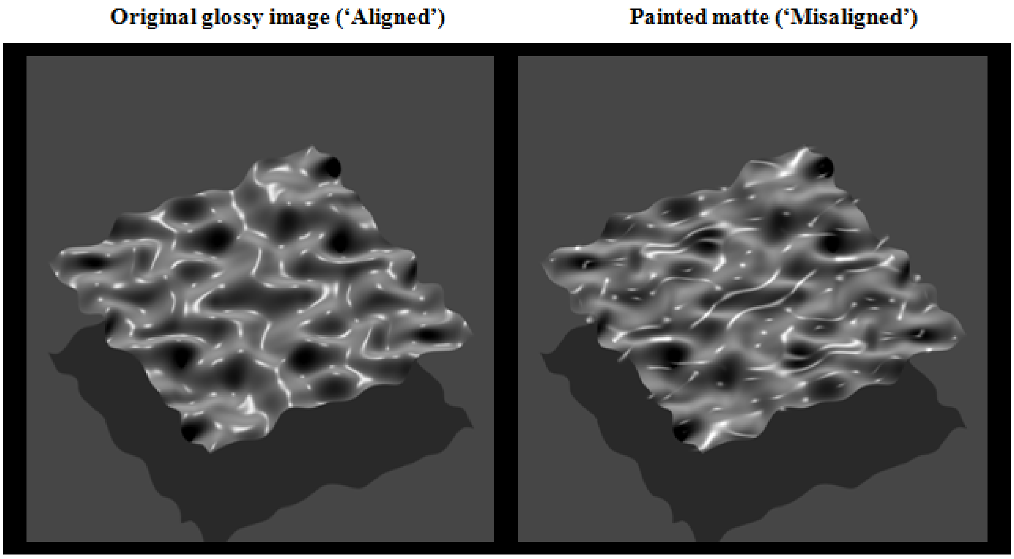

Figure 12. Stimuli used in Experiment 1. The image of a planar surface where the specular highlight layer and diffuse shading profile are aligned with each other (Left) When those two components are not aligned, the surface looks like 'painted matte' (Right).

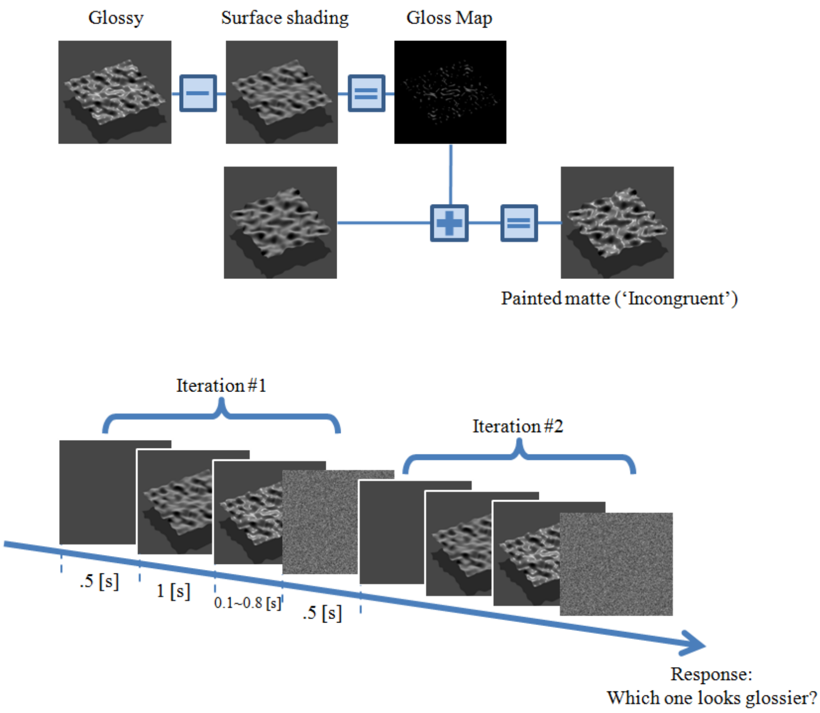

Figure 13. Schematic diagram for generating stimuli and experimental procedure from Experiment 1. The method to generate misaligned surface images is shown. The gloss map was extracted from the original glossy image, and superimposed on the other matte surface image (Top). The procedure for the behavioral experiment (Bottom)

$$y = f(x, z) = \sum_{k=1}^{15} \sin(\pi a_k[x/\lambda + \pi] + \pi b_k[z/\lambda + \pi]). \quad \text{(Equation 1)}$$

$$I = k_a + (1-a)k_d \cos(\theta) + ak_s \cos^n(\alpha), \quad \text{(Equation 2)}$$

At each pixel location, the inner product of the surface's normal and illumination directions (Lambertian shading) was calculated; the reflectance was also calculated based on the Ward lighting model, and the underlying equation is shown in Equation 2. In Blender, the specular parameter was set to .7. The location of a directional light source was placed in the scene at the coordinates (1, 1, .7) in x-, y-, and z-planes respectively. The polar angle between a surface and a light source was 55 degrees. The viewer-surface angle was set to 30 degrees. The maximum luminance of the monitor was 85.0 cd/m^2. The background was set to relative RGB = (0, 0, 30), which appeared in the blue color. The cast shadows were rendered in black (8 cd/m^2) and a planar surface (8 by 8 degrees in visual angle) was rotated 30 degrees around the Z-axes in order to enhance the depth impression of space. The fixation red dot was presented in the center of the stimulus all the time. The examples of stimuli used in the experiment are shown in Figure 12.

<u>Pillar objects</u>

A pillar-type 3D object was generated based on a planar surface. First, planar surfaces were generated by the same procedure in Experiment 1, and those surfaces were rolled up around a cylinder-shaped polygon mesh. The circumference of the cylinder was matched to the length of the side of a planar surface. After generating a volumetric object, we used Blender 3D software to render objects in 3D space. All material properties and light sources were the same as in Experiment 1. To generate a movie clip of a rotating pillar, the pillar was rotated 1 degree for each frame and rendered separately (Figure 14).

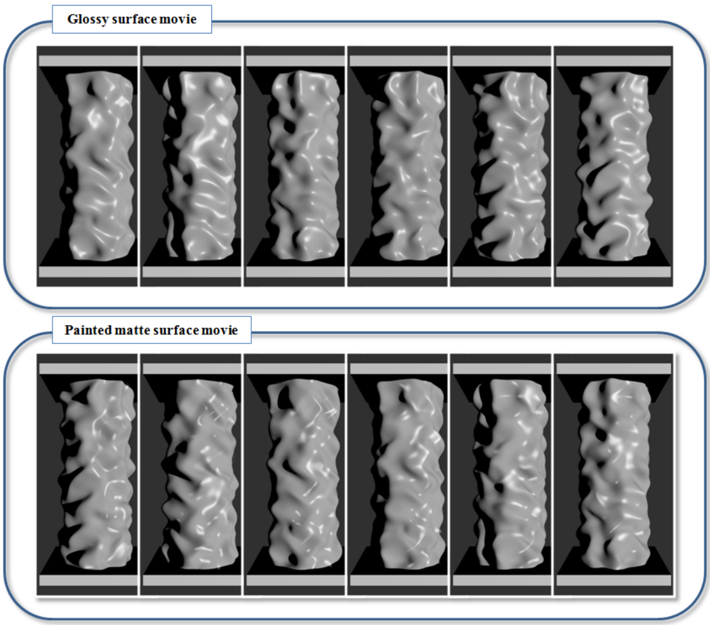

Figure 14. The movie example used in Experiment 2. The example of six frames from a rotating pillar movie clip was shown. The specular highlights of the rotating pillar are aligned to its surface shape, appearing glossy (Top). Movie frames of a misaligned rotating pillar were shown. For the overall frames, each surface image does not look glossy, but rather, like painted matte (Bottom).

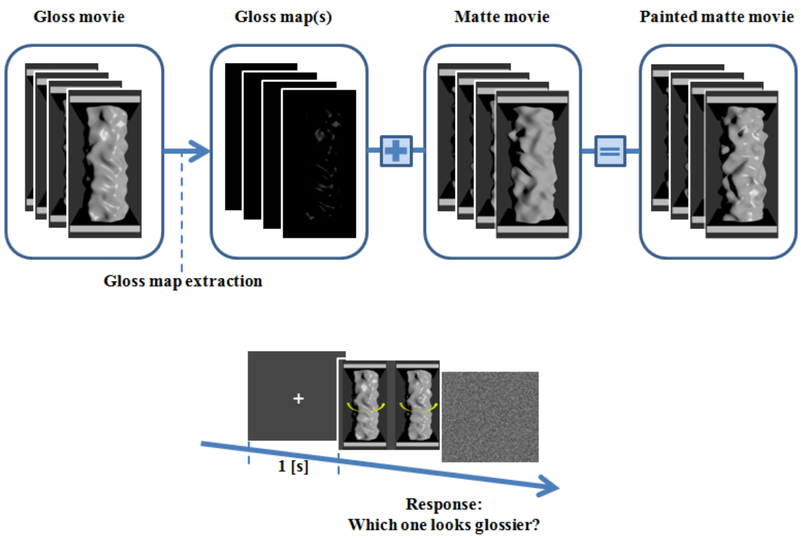

Figure 15. Schematic diagram for generating stimuli and experimental procedure from Experiment 2. The figure Shows the procedure of how a gloss map was extracted in order to generate aligned and misaligned pillar movie clips (Top). The procedure of behavioral experiment (Bottom)

Gloss map manipulation

Each surface was rendered with two reflectance parameters, one at full reflectance (alpha = .7), and the other at complete matte (alpha = 0). We subtracted the pixel intensity map of the matte surface from the glossy surface, which produced the 'gloss map' in Figure 15. Then, the gloss map was superimposed on a randomly chosen matte image from the image data set. The overall luminance and contrast (RMS) were matched across all planar images generated from the procedure.

Procedure

In Experiment 1, simulated aligned and misaligned planar surfaces were presented based on two interval forced choice paradigms (2-IFC), as shown in Figure 13 (Bottom). Before the beginning of each trial, observers viewed a fixation for .5 [s], followed by a matte surface image that was presented for 1 [s]. After than, one of the aligned or misaligned specular highlight layers was superimposed on the matte surface and remained for between .1 and .8 [s] depending on the experimental condition. The noise masking image was followed for 500 [ms] by the end of each interval. During each trial, observers were asked to maintain a fixation on the center dot all the time; when the blank screen appeared, observers were asked to report which of either the first or second surface looked glossier than the other by pressing a predefined button. The stimulus duration of each trial randomly varied based on three levels of a time length (.1 [s], .4 [s], and .8 [s]). Observers completed 120 trials for each viewing time level. The sequence of aligned and misaligned surface presentation within each trial was pseudo-randomized and the frequency of appearance was counter-balanced within one session.

In Experiment 2, before the beginning of each trial observers viewed a fixation for 1 [s], and then two rotating pillar movies were presented at the same time, side-by-side on a monitor (Figure 15, Bottom). One movie showed an aligned pillar, and the other showed a misaligned specular component. Observers performed 2 alternative forced choice tasks: they were asked to report which of two pillars looked glossier than the other by pressing a left- and right-arrow button on a keyboard. The pillar stimuli were continuously rotating until the observer's input finished. For each trial, the movies had four different speed levels (.3, .8, 1.3, and 1.8 [revolution/second]); the speed level was randomly chosen, and the sequence was pseudo-randomly intermixed and counter-balanced. The two pillars presented side-by-side had the same rotation speed. Observers completed a total of 80 trials for each speed level.

4.3 Results and discussion

Experiment 1: Glossiness under different lengths of viewing time

The goal of Experiment 1 was to measure the discrimination accuracy of a target stimulus under various viewing time conditions. The target stimulus was made of a specular highlight layer that was aligned to surface shading ('physically correct'), and a comparison stimulus was made of misaligned ('physically incorrect') specular highlight layers. We calculated the probabilities of 'aligned' target images being selected as glossier than the other comparatively 'misaligned' images. Figure 16 (Top) shows the probability of a target to be selected as a function of the amount of viewing time.

The accuracy was significantly influenced by the viewing time, $F(2,39) = 16.38$, $p < 0.0001$. For the longest time condition (.8 [sec]), the probability of being glossier was higher than that of the intermediate exposure time (.4 [sec]), $T(14) = 1.8$, $p < .05$. Similarly, the probability of the intermediate time condition was higher than that of the shortest time condition (.1 [sec]), $T(14) = 4.53$, $p < 0.001$. As indicated by the linear regression (Figure 3, Bottom), the probability showed a clear ascending trend when more time was given to observers (Beta = .264, $p < .0005$).

The results shows that the probabilities of target stimuli selected as being glossier declined when the length of exposure time shortened. It should be noted that a physical reflectance of a target stimuli did not vary, and consequently, the overall luminance and contrast of stimuli remained the same.

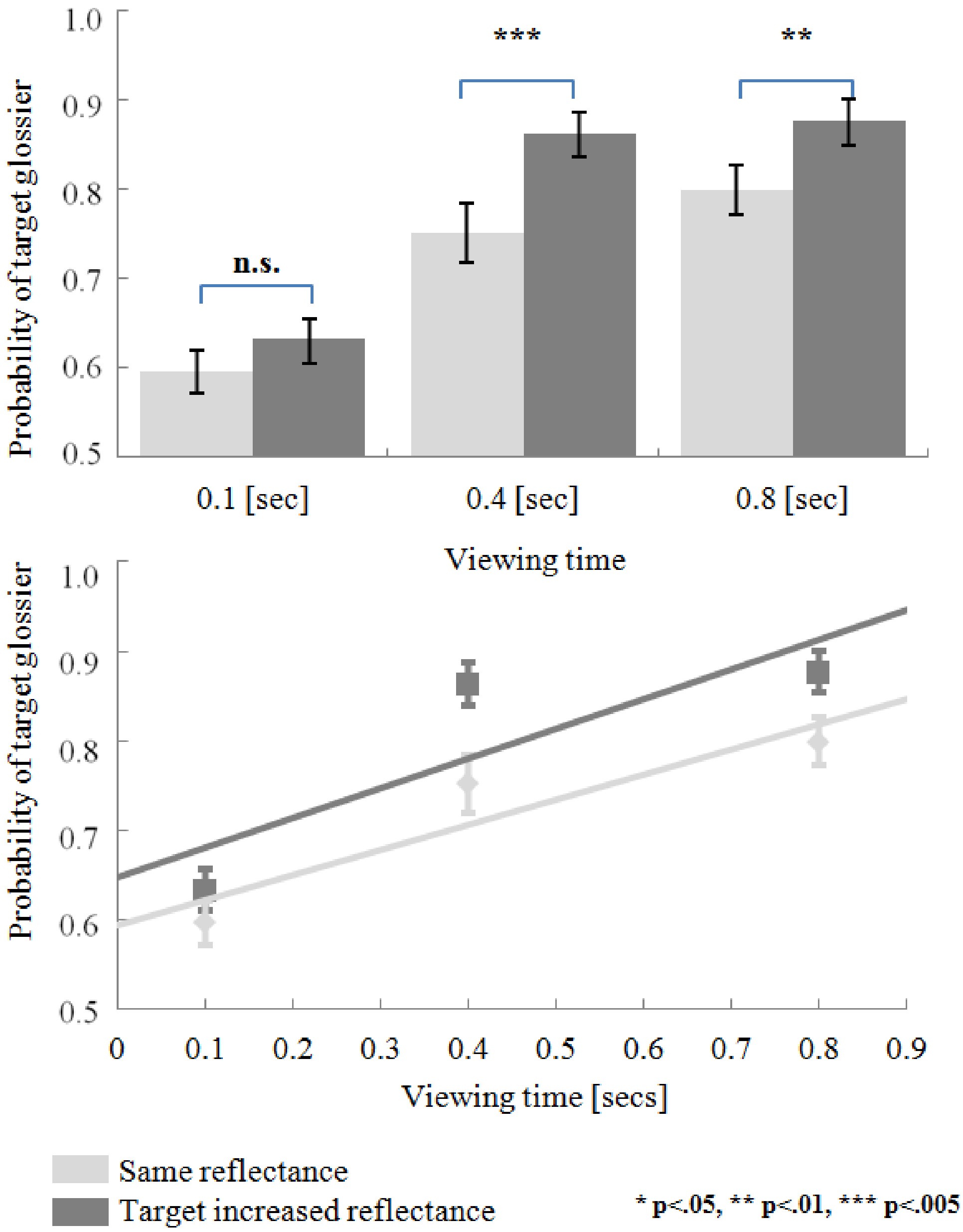

Figure 16. Results from Experiment 1. The x-axis indicates the different level of viewing time, and the y-axis indicates the probability of a target to be selected as glossier than a baseline, 50% (Top). Linear regression plotted based on observer's responses (Bottom).

The results indicate that the perceived glossiness of a target surface is negatively influenced by viewing time because the probability of being selected as glossier was reduced. It is notable that the perceived glossiness reduced rapidly in 100 [ms] viewing time conditions, which suggests that the time of exposure did not linearly influence the discrimination performance of glossiness.

In order to test whether the decrease of probability would be due to the poor encoding of luminance and the contrast intensity of a target stimulus, we increased the reflectance parameter of target stimuli and rendered them again. The results were shown in Figure 16 (Bottom) as dark-gray colored bars and lines. The viewing time influenced the perceived glossiness of the high reflectance version of surfaces, $F(2,39) = 33.24$, $p < 0.0001$, which was very similar to the result obtained in Experiment 1. The perceived glossiness of a target stimulus in the shortest time condition was lower than in the other two conditions ($\mu = .86$ for the intermediate and $\mu = .87$ for the longest time condition). Although high reflectance conditions shows that the increase of physical reflectance can enhance the discrimination ability, this enhancement effect was only observed when the viewing time was longer than 400 [ms]. That is, the effect of elevating physical reflectance diminished in the shortest viewing time, which is 100 [ms].

Experiment 2: The influence of object motion on perceived glossiness

In previous experiments, a limited amount of viewing time led to a reduced probability for a target surface to be selected as glossier. These results indicate that the target surface looked less glossy when the observers' exposure times to stimuli were shortened. However, this phenomenon may simply reflect the poor encoding of visual input due to the lack of attending time to stimuli.

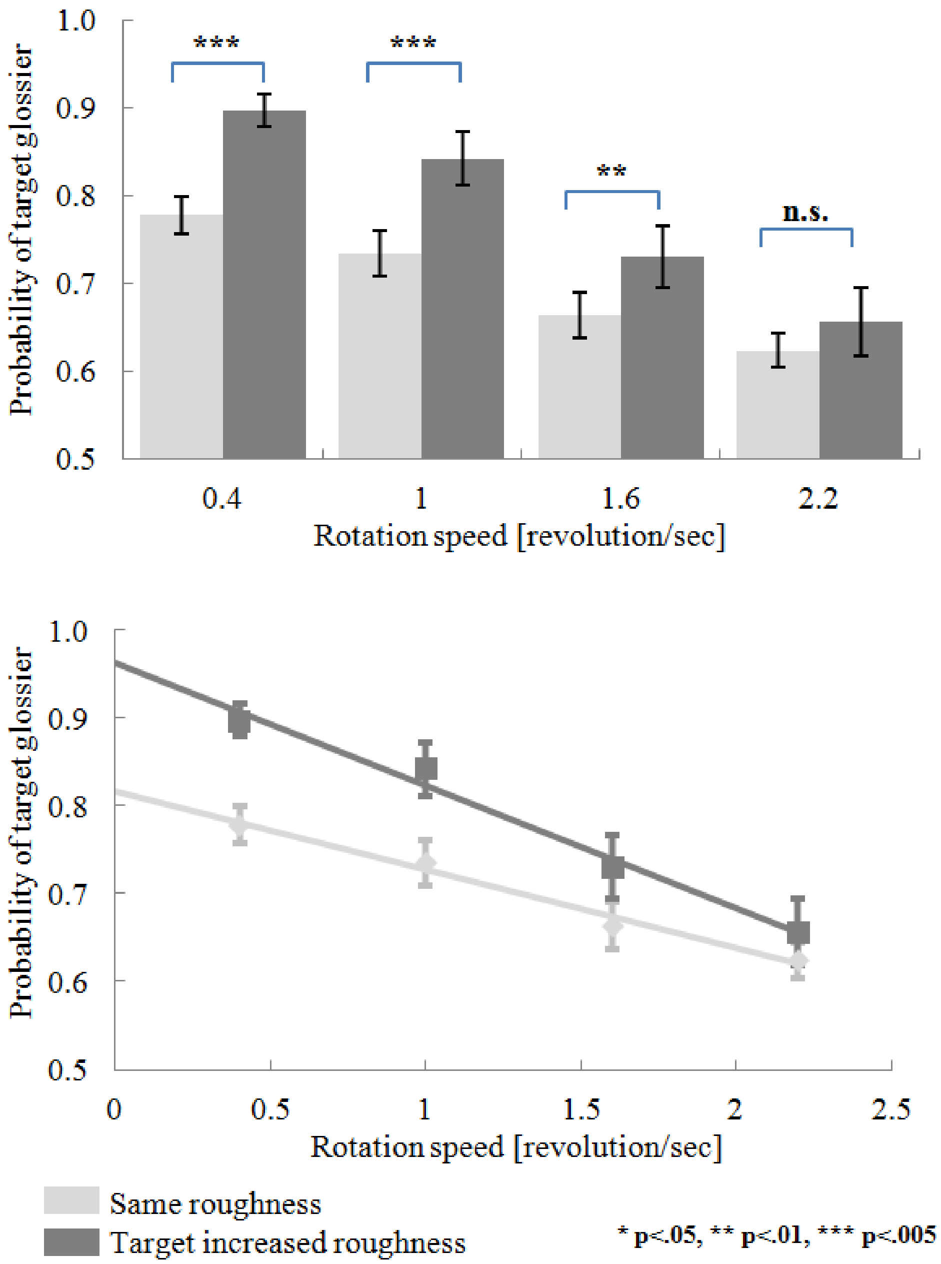

Figure 17. Results from Experiment 2. The x-axis indicates the different levels of a rotation speed in the number of revolutions per seconds, and the y-axis indicates the probability of a target (aligned pillar) to be selected as glossier more than the baseline, 50% (Top). Linear regression plotted based on observer's responses (Bottom).

In Experiment 2, we added a motion property of aligned and misaligned highlight objects. The goal of this experiment was to examine whether a limited viewing time for each frame in continuously rotating motion could influence the overall glossiness. One major distinction between Experiment 1 and 2 is that unlike the static surface in Experiment 1, observers had enough time to fixate on the images in Experiment 2. Because observers were not disrupted in attending to stimuli for a long time, we could rule out the possibility that lack of attention would affect perceived surface glossiness. Figure 17 shows the probability of a target rotating pillar to be selected as glossier than a comparison rotating pillar. The results showed that the rotation speed of pillar objects significantly influenced the perceived glossiness of a target stimulus ($F(3,59) = 8.72$, $p < .0001$). The probability of a target stimulus being selected as glossier was 28% more than the baseline (50%) when objects were rotating at the slowest speed (.4 [rev/sec]), whereas the probability was only 12% more in the fastest rotating condition (2.2 [rev/sec]), with an overall probability of 62% (the baseline is 50%) in Figure 17 (Top). The two intermediate speed conditions (1 and 1.6 rev/sec) also showed that the slower speed condition (1 rev/sec) yielded more probability. Although perceived glossiness reduces as rotating speed gets faster, the probability of the fastest speed condition (2.2 rev/sec) was still above chance level (50%) in Figure 17 (Bottom)

To test whether surface roughness could enhance the perceived glossiness of a rotating pillar object, we added more curvatures to both the target and comparison pillars. The results are depicted in Figure 17 with dark-gray colored bars and lines. The overall trend was not much different from the original result (Experiment 2): the probabilities decreased when rotation speed got faster. The probabilities in the slowest condition were 48% higher than the baseline (50%) in the slowest speed condition, and 15% more in the fastest speed condition. The probability difference between the slowest and fastest condition was significant, $T(14) = 5.99$, $p<.0001$.

The overall probability decrease clearly demonstrates there is a negative relationship between probabilities of being selected and rotation speed, which is very similar to the results obtained in Experiment 1. Both Experiment 1 and 2 allow us to reach one conclusion: the time spent viewing specular objects influences their perceived glossiness, even though their physical reflectance was invariant. Together with many previous studies showing that perceiving optical properties of an object could be influenced by many contextual factors, this study suggests that viewing time is also one important factor to determining the perceived glossiness of an object in an image. In addition, the results showed that surface roughness can enhance the discrimination sensitivity to glossiness presumably due to that complex curvature aided the visual system in computing the spatial correspondence between highlights and surface geometry.

4.4 General discussion

In the present study, we investigated the temporal processing of surface glossiness. The results showed that the viewing time of a surface can significantly modulate the degree of perceived glossiness. In Experiment 1, the limited amount of time viewing a static surface negatively influenced the discrimination accuracy of a physically correct target surface. When less viewing time was given, the probability of selecting a target stimulus as being glossier was reduced. In Experiment 2, when we tested the influence of rotating speed of a 3D pillar object, the speed also negatively influenced the discrimination accuracy. The fast speed disrupts the visual system’s ability to assess surface glossiness, resulting in reduced sensitivity to discriminate the glossy surface from the painted matte surface. These results suggest that viewing time is one crucial factor to properly estimate the surface reflectance, since the visual system needs time to compute the spatial correspondence between luminance maxima (specular highlights) and surface geometry.

One may claim that some image features, such as a shape or elongated direction of specular highlights, can influence discrimination accuracy. For example, observers could rely on certain low-level image cues to detect the target surface, rather than comparing perceived glossiness between a target and a comparison stimulus. Indeed, elongated or stretched out appearance of highlights is closely related to three-dimensional surface geometry (Fleming et al., 2004), and the perceived size and contrast of highlights highly correlated with the degree of perceived gloss (Marlow & Anderson, 2013).

Because we used misaligned specular highlights, most likely the appearance of highlights could be unnatural in their size or elongated directions. If this is true, then observers could use these appearances as a cue to perform a task. However, specular highlight layers were not randomly generated in our experiment. Rather, we obtained those layers from different stimuli sets, and thus there were no 'unnatural' specular highlights. Both congruent and incongruent surface images were seemingly not much different in terms of their highlights. Furthermore, both images had very similar image statistics, such as mean luminance and contrast or shape of luminance histogram (skewness); thus it is not possible to discriminate the aligned target stimulus from the misaligned comparison stimulus based on local image features or global configurations of an image. Recent studies suggest that the visual system can rapidly calculate low-level image statistics of objects and scene. For instance, the mean size of a set of shapes (Ariely, 2001; Chong & Treisman, 2005), the average orientation of a pattern (Parkes et al., 2001), or the texture statistics of a scene (Renninger & Malik, 2004) can be processed within about 100 [ms]. In our experiment, however, observers needed at least 300 ms of viewing time to reach 75% discrimination accuracy. Once again, the results of our experiments reflect that perceived glossiness was influenced by viewing time for observers, and processing surface glossiness could be limited depending on a lack of viewing time. This is because the visual system relies on the computation of the spatial relationship between highlights and surface geometry, rather than just detecting low-level image features to assess surface glossiness.

Previously, it has been shown that the spatial relationship ('spatial congruence') between diffuse shading profiles (surface geometry) and the location of specular highlights is a crucial factor in evoking the gloss sensation. In a series of experiments conducted by Anderson and his colleagues, the rotated or translated specular layers of a surface ("'gloss map") resulted in a monotonic decrease of perceived glossiness. Furthermore, this relationship was formulated as a function of the amount of deviation from rotational and translational highlights to its original location. Although the data was not shown here, our pilot experiments also showed similar results; observers were asked to score the perceived glossiness of congruent and incongruent specular images, and the results showed that images which were based on spatially congruent diffuse and specular components obtained higher degrees of glossiness than spatially incongruent image sets. Thus spatial correspondence is very important information that the visual system relies on to judge surface gloss.

In Experiment 2, our study demonstrated that not only viewing time of a surface, but also the speed of an object can influence the perceived glossiness. When objects were rotating fast (2 rev/sec), observers could not distinguish an aligned target pillar from an unaligned comparison pillar, meaning that their perceived glossiness was indistinguishable. The image of a moving object continuously changes its visual appearance. For a fixed amount of time, more information changes when an object moves faster. The visual system requires assessing glossiness more quickly during a limited amount of time. Previous studies showed that retinal-image motion produces characteristic optic flow patterns, and these patterns could generally enhance the glossiness of the object (Doerschner et al., 2011; Tani et al., 2013; Sakano & Ando, 2010). These results show that the visual system is able to extract and process motion information, showing better performance in estimating surface gloss. Our results in Experiment 2 showed a change in performance of estimating glossiness depending on the speed of an object. Interestingly, we found a temporal limitation to discriminate gloss from the local luminance maxima of the misaligned pillar. The discrimination accuracy of a target stimulus from a comparison stimulus declined as rotation speed sped up. These results suggest two possibilities: One is that the perceived glossiness of a target surface image reduced due to fast image changes. However, as noted earlier, object motion promotes a general glossiness, providing more image information of surface material. Alternatively, the visual system may lose its spatial sensitivity in assessing the relationship between surface shape and highlights. In this case, both stimuli seem to exhibit similar glossiness, resulting in reduced discrimination accuracy. Indeed, most observers experienced both stimuli as looking evenly glossier when the stimuli were rotating at the maximum speed (2.2 rev/sec), although their reports were anecdotal. Taken together the results obtained in Experiment 1 and 2 suggest that glossiness perception is closely related to the temporal processing of surface information.

Specifically, the visual system has a limited temporal processing capability for evaluating the spatial relationship between surface shape and highlight. The fast motion of an object can disrupt the computational process of surface information, resulting in sensitivity loss in assessing the physical correctness of image information on a surface.

4.5 Conclusion

We tested whether the glossiness of a surface could be influenced by viewing time of the surface. The results indicate that 1) the probability of a static planar surface to be selected as glossier reduced when viewing time was limited, and 2) the probability of a rotating pillar to be selected as glossier also reduced when the objects were rotating very fast. Both results suggest that the visual system has a limited temporal processing ability to extract surface information and estimate surface reflectance. Unlike object/scene perceptions, in order to perceive a gloss, the visual system needs a longer time than 300 [ms]. The findings imply that the underlying mechanism of glossiness perception differs from the mechanism underlying object/scene perception. Presumably, perceiving glossiness is a process to compute the spatial relationship between surface shapes and bright spots. The time required to perceive gloss reflects that visual computation processes take longer than shape information processing in recognizing an object.

Chapter 5: Concluding remarks

The visual system extracts basic image features from sensory input. It also interprets the relationship between those features, integrating them into more complex object information, which is called the "constructive process". In the chapter 2, we discussed whether this process could function without the observer's awareness. We found that the viewpoint of an invisible cube could evoke an adaptation aftereffect. These results suggest that the visual system is able to construct 3D structure information even in the absence of awareness. Another crucial role of the visual system is constructing a single, coherent representation of a visual object by resolving perceptual ambiguity. We discussed the underlying mechanism that resolves this ambiguity in the chapter 3. Our specific research question was whether or not different bistable perceptions have the same mechanism for resolving ambiguity. We observed that the temporal dynamics of perceptual alternation were different depending on the types of ambiguous figures, the presented location in the visual field, and the observer's handedness. The results suggest that ambiguous sensory input is processed by multiple neural substrates, rather than a single common substrate. For example, the alternation in binocular rivalry is modulated by both the eyes and hemispheric factors, whereas the alternation in the Necker cube is most likely modulated by hemisphere-level mechanisms. Finally, in the chapter 4, we measured how much the duration of viewing time and the speed of a moving object influence the visual system's ability to process more complex surface properties (material information). When a shorter time was given to view an object's surface, it was more likely that the visual system inaccurately estimated the surface reflectance. A similar tendency was also observed when an object was rotating very quickly. Both results clearly indicated that perceived glossiness varied depending on how long the stimulus was exposed to the visual system. We argue that in order to properly estimate complex surface information, such as glossiness, the visual system needs a minimum amount of processing time.

As an intermediate level of processing in the visual hierarchy, the "constructive process" can operate even when observers are not aware of sensory input. The neural substrates underlying this process are presumably situated beyond the stage in which binocular information converges. This is because a certain bistable perception (e.g., the Necker cube) shows that temporal dynamics in both eyes have very similar patterns. A narrow range of implications in our study would be that 3D structures and viewpoints of an object can be processed without awareness. For broader implications, the outcome of unconsciously processed sensory information could influence the observer's behavioral responses related to visual information processing. For example, the present study can account for how visually guided viewer-to-object interactions (e.g., reaching and grasping) could arise without observer's explicit awareness. Indeed, the observational study of blindsight patients showed that in the absence of visual awareness, patients could execute appropriate visuo-motor actions. Future research may address the extent of this unconscious processing. 3D structure and viewpoint are just two types of object information; therefore, it still remains to be answered what other kinds of object features can be processed in the absence of awareness.

Although the constructive process can operate without an observer's explicit awareness, this does not mean that the process occurs instantaneously or simultaneous to image information entering the visual system. To interpret more complex object features, such as reflectance or the geometry of an object's surface, the visual system needs a minimum amount of time. When an object is moving very fast or briefly presents (~300 ms), the visual system more likely fails to integrate features. As a result, the percept of a scene would be fragmentary or incomplete. Specifically, a series of experiments in chapter 4 indicates that the viewing time and moving speed of an object influence estimation of its surface glossiness. The results imply that the constructive process presumably consists of computations in neural substrates, and the accuracy of computation is proportional to the observer's viewing time. As another major implication, the present idea can account for the extent to which information can or cannot be processed when rapid eye movement occurs (saccade). In this case, the present study predicts that complex surface properties, such as the reflectance and glossiness of an object, would not be accurately recognized due to the lack of viewing time. In the future, one can investigate how processing time relates to other types of object features, such as the complexity of the object's shape, size, and familiarity.

References

Alais, D., O'Shea, R. P., Mesana-Alais, C., & Wilson, I. G. (2000). On binocular alternation. *Perception*, 29, 1437–1445

Alexander, L. T., & Bricker, P. D. (1952). Figure-ground contrast and binocular rivalry. *Journal of experimental psychology*, *44*(6), 452.

Almeida, J., Mahon, B. Z., & Caramazza, A. (2010). The role of the dorsal visual processing stream in tool identification. *Psychological science*, *21*(6), 772-778.

Almeida, J., Mahon, B. Z., Nakayama, K., & Caramazza, A. (2008). Unconscious processing dissociates along categorical lines. *Proceedings of the National Academy of Sciences*, *105*(39), 15214-15218.

Anderson, B. L., & Kim, J. (2009). Image statistics do not explain the perception of gloss and lightness. *Journal of vision*, *9*(11), 10.

Ariely, D. (2001). Seeing sets: Representation by statistical properties. *Psychological Science*, *12*(2), 157-162.

Beck, J., & Prazdny, S. (1981). Highlights and the perception of glossiness. *Attention, Perception, & Psychophysics*, *30*(4), 407-410.

Bentin, S., Allison, T., Puce, A., Perez, E., & McCarthy, G. (1996). Electrophysiological studies of face perception in humans. *Journal of cognitive neuroscience*, *8*(6), 551-565.

Berardi, N., & Fiorentini, A. (1991). Visual field asymmetries in pattern discrimination: A sign of asymmetry in cortical visual field representation?. *Vision research*, *31*(10), 1831-1836.

Blake, R. (2001). A primer on binocular rivalry, including current controversies. *Brain and mind*, *2*(1), 5-38.

Blake, A., & Bülthoff, H. (1990). Does the brain know the physics of specular reflection?. *Nature*, *343*(6254), 165-168.

Blake, R., & Fox, R. (1974). Adaptation to invisible gratings and the site of binocular rivalry suppression.

Blake, R., & Logothetis, N. K. (2002). Visual competition. *Nature Reviews Neuroscience*, *3*(1), 13-21.

Blakemore, C. T., & Campbell, F. W. (1969). On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images. *The Journal of physiology*, *203*(1), 237-260.

Boring, E. G. (1942). *Sensation and perception in the history of experimental psychology*. D. Appleton-Century Company.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial vision*, *10*, 433-436.

Breese, B. B. (1909). Binocular rivalry. *Psychological Review*, *16*(6), 410.

Borod, J. C., St Clair, J., Koff, E., & Alpert, M. (1990). Perceiver and poser asymmetries in processing facial emotion. *Brain and cognition*, *13*(2), 167-177.

Bryden, M. P., Ley, R. G., & Sugarman, J. H. (1982). A left-ear advantage for identifying the emotional quality of tonal sequences. *Neuropsychologia*, *20*(1), 83-87.

Carlson, V. R. (1953). Satiation in a reversible perspective figure. *Journal of Experimental Psychology*, *45*(6), 442.

Cass, J., Johnson, A., Bex, P. J., & Alais, D. (2012). Orientation-specificity of adaptation: isotropic adaptation is purely monocular. *PloS one*, *7*(11), e47425.

Cave, C. B., Blake, R., & McNamara, T. P. (1998). Binocular rivalry disrupts visual priming. *Psychological Science*, *9*(4), 299-302.

Chen, X., & He, S. (2003). Temporal characteristics of binocular rivalry: visual field asymmetries. *Vision research*, *43*(21), 2207-2212.

Chong, S. C., & Treisman, A. (2005). Statistical processing: Computing the average size in perceptual groups. *Vision research*, *45*(7), 891-900.

Corballis, M. C., & Sergent, J. (1989). Hemispheric specialization for mental rotation. *Cortex*, *25*(1), 15-25.

Costello, P., Jiang, Y., Baartman, B., McGlennen, K., & He, S. (2009). Semantic and subword priming during binocular suppression. *Consciousness and cognition*, *18*(2), 375-382.

Curcio, C. A., Sloan, K. R., Packer, O., Hendrickson, A. E., & Kalina, R. E. (1987). Distribution of cones in human and monkey retina: individual variability and radial asymmetry. *Science*, *236*(4801), 579-582.

Davidoff, J. B. (1977). Hemispheric differences in dot detection. *Cortex*, *13*(4), 434-444.

Doerschner, K., Boyaci, H., & Maloney, L. T. (2010). Estimating the glossiness transfer function induced by illumination change and testing its transitivity. *Journal of Vision*, *10*(4), 8.

Doerschner, K., Fleming, R. W., Yilmaz, O., Schrater, P. R., Hartung, B., & Kersten, D. (2011). Visual motion and the perception of surface material. *Current Biology*, *21*(23), 2010-2016.

Downing, P. E., Jiang, Y., Shuman, M., & Kanwisher, N. (2001). A cortical area selective for visual processing of the human body. *Science*, *293*(5539), 2470-2473.

Evans, K. K., & Treisman, A. (2005). Perception of objects in natural scenes: is it really attention free?. *Journal of Experimental Psychology: Human Perception and Performance*, *31*(6), 1476.

Fang, F., & He, S. (2005). Cortical responses to invisible objects in the human dorsal and ventral pathways. *Nature neuroscience*, *8*(10), 1380-1385.

Fahle, M. (1987). Naso-temporal asymmetry of binocular inhibition. *Investigative Ophthalmology & Visual Science*, 28(6), 1016–1017.

Fahle, M., & Schmid, M. (1988). Naso-temporal asymmetry of visual perception and of the visual cortex. *Vision Research*, 28(2), 293–300.

Falconbridge, M., Ware, A., & MacLeod, D. I. (2010). Imperceptibly rapid contrast modulations processed in cortex: evidence from psychophysics. *Journal of vision*, *10*(8), 21.

Fei-Fei, L., Iyer, A., Koch, C., & Perona, P. (2007). What do we perceive in a glance of a real-world scene?. *Journal of vision*, *7*(1), 10.

Fleming, R. W., Dror, R. O., & Adelson, E. H. (2003). Real-world illumination and the perception of surface reflectance properties. *Journal of Vision*, *3*(5), 3.

Fleming, R. W., Torralba, A., & Adelson, E. H. (2004). Specular reflections and the perception of shape. *Journal of Vision*, *4*(9), 10.

Freeman, A. W. (2005). Multistage model for binocular rivalry. *Journal of Neurophysiology*, *94*(6), 4412-4420.

Galaburda, A. M., Rosen, G. D., & Sherman, G. F. (1990). Individual variability in cortical organization: its relationship to brain laterality and implications to function. *Neuropsychologia*, *28*(6), 529-546.

Gilbert, C., & Bakan, P. (1973). Visual asymmetry in perception of faces. *Neuropsychologia*, *11*(3), 355-362.

Goebel, R., Khorram-Sefat, D., Muckli, L., Hacker, H., & Singer, W. (1998). The constructive nature of vision: direct evidence from functional magnetic resonance imaging studies of apparent motion and motion imagery. *European Journal of Neuroscience*, *10*(5), 1563-1573.

Greene, M. R., & Oliva, A. (2009). The briefest of glances: The time course of natural scene understanding. *Psychological Science*, *20*(4), 464-472.

Günther, J., Chen, T., Goesele, M., Wald, I., & Seidel, H. P. (2005, November). Efficient acquisition and realistic rendering of car paint. In *Vision, Modeling, and Visualization* (Vol. 5, pp. 487-494).

Hamilton, C. R., & Vermeire, B. A. (1988). Complementary hemispheric specialization in monkeys. *Science*, *242*(4886), 1691-1694.

Harris, J. P. (1980). How does adaptation to disparity affect the perception of reversible figures?. *The American journal of psychology*, 445-457.

Harris, J. J., Schwarzkopf, D. S., Song, C., Bahrami, B., & Rees, G. (2011). Contextual illusions reveal the limit of unconscious visual processing. *Psychological science*, *22*(3), 399-405.

Heilman, K. M., & Van Den Abell, T. (1980). Right hemisphere dominance for attention: The mechanism underlying hemispheric asymmetries of inattention (neglect). *Neurology*, *30*(3), 327-327.

Ho, Y. X., Landy, M. S., & Maloney, L. T. (2008). Conjoint measurement of gloss and surface texture. *Psychological Science*, *19*(2), 196-204.

Hoptman, M. J., & Levy, J. (1988). Perceptual asymmetries in left-and right-handers for cartoon and real faces. *Brain and cognition*, *8*(2), 178-188.

Howard, I. P. (1961). An investigation of a satiation process in the reversible perspective of revolving skeletal shapes. *Quarterly Journal of Experimental Psychology*, *13*(1), 19-33.

Hutton, J. T., & Palet, J. (1986). Lateral saccadic latencies and handedness. *Neuropsychologia*, *24*(3), 449-451.

Iachini, T., Iavarone, A., Senese, V. P., Ruotolo, F., & Ruggiero, G. (2009). Visuospatial memory in healthy elderly, AD and MCI: a review. *Current aging science*, *2*(1), 43-59.

Intraub, H. (1981). Rapid conceptual identification of sequentially presented pictures. *Journal of Experimental Psychology: Human Perception and Performance*, *7*(3), 604.

Jiang, Y., Costello, P., & He, S. (2007). Processing of invisible stimuli: Advantage of upright faces and recognizable words in overcoming interocular suppression. *Psychological science*, *18*(4), 349-355.

Kang, M. S., Blake, R., & Woodman, G. F. (2011). Semantic analysis does not occur in the absence of awareness induced by interocular suppression. *The Journal of Neuroscience*, *31*(38), 13535-13545.

Kanwisher, N., & Yovel, G. (2006). The fusiform face area: a cortical region specialized for the perception of faces. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *361*(1476), 2109-2128.

Kaushall, P. (1975). Functional asymmetries of the human visual system as revealed by binocular rivalry and binocular brightness matching. *American Journal of Optometry & Physiological Optics*, 52(8), 509–520.

Kovács, I., Papathomas, T. V., Yang, M., & Fehér, Á. (1996). When the brain changes its mind: Interocular grouping during binocular rivalry. *Proceedings of the National Academy of Sciences of the United States of America*, 93, 15508–15511.

Knecht, S., Dräger, B., Deppe, M., Bobe, L., Lohmann, H., Flöel, A., ... & Henningsen, H. (2000). Handedness and hemispheric language dominance in healthy humans. *Brain*, *123*(12), 2512-2518.

Lehmkuhle, S. W., & Fox, R. (1975). Effect of binocular rivalry suppression on the motion aftereffect. *Vision research*, *15*(7), 855-859.

Leopold, D. A., & Logothetis, N. K. (1999). Multistable phenomena: changing views in perception. *Trends in cognitive sciences*, *3*(7), 254-264.

Levelt, W. J. (1967). Note on the distribution of dominance times in binocular rivalry. *British Journal of Psychology*, *58*(1-2), 143-145.

Levy, J., Heller, W., Banich, M. T., & Burton, L. A. (1983). Asymmetry of perception in free viewing of chimeric faces. *Brain and cognition*, *2*(4), 404-419.

Long, G. M., Toppino, T. C., & Mondin, G. W. (1992). Prime time: Fatigue and set effects in the perception of reversible figures. *Perception & psychophysics*, *52*(6), 609-616.

Luh, K. E., Redl, J., & Levy, J. (1994). Left-and right-handers see people differently: Free-vision perceptual asymmetries for chimeric stimuli. *Brain and cognition*, *25*(2), 141-160.

Marlow, P. J., & Anderson, B. L. (2013). Generative constraints on image cues for perceived gloss. *Journal of vision*, *13*(14), 2.

Maruya, K., Watanabe, H., & Watanabe, M. (2008). Adaptation to invisible motion results in low-level but not high-level aftereffects. *Journal of Vision*, *8*(11), 7.

Montoro, P. R., Luna, D., & Ortells, J. J. (2014). Subliminal Gestalt grouping: Evidence of perceptual grouping by proximity and similarity in absence of conscious perception. *Consciousness and cognition*, *25*, 1-8.

Moradi, F., Koch, C., & Shimojo, S. (2005). Face adaptation depends on seeing the face. *Neuron*, *45*(1), 169-175.

Motoyoshi, I., Nishida, S. Y., Sharan, L., & Adelson, E. H. (2007). Image statistics and the perception of surface qualities. *Nature*, *447*(7141), 206-209.

Nicholls, M. E. (1994). Hemispheric asymmetries for temporal resolution: A signal detection analysis of threshold and bias. *The Quarterly Journal of Experimental Psychology*, *47*(2), 291-310.

Nicodemus, F. E. (1965). Directional reflectance and emissivity of an opaque surface. *Applied optics*, *4*(7), 767-773.

Nishida, S. Y., Ashida, H., & Sato, T. (1994). Complete interocular transfer of motion aftereffect with flickering test. *Vision research*, *34*(20), 2707-2716.

Nishida, S. Y., & Shinya, M. (1998). Use of image-based information in judgments of surface-reflectance properties. *JOSA A*, *15*(12), 2951-2965.

Obein, G., Knoblauch, K., & Viéot, F. (2004). Difference scaling of gloss: Nonlinearity, binocularity, and constancy. *Journal of Vision*, *4*(9), 4.

Okubo, M., & Nicholls, M. E. (2008). Hemispheric asymmetries for temporal information processing: Transient detection versus sustained monitoring. *Brain and cognition*, *66*(2), 168-175.

Oliva, A., & Torralba, A. (2001). Modeling the shape of the scene: A holistic representation of the spatial envelope. *International journal of computer vision*, *42*(3), 145-175.

Olkkonen, M., & Brainard, D. H. (2010). Perceived glossiness and lightness under real-world illumination. *Journal of vision*, *10*(9), 5.

Oren, M., & Nayar, S. K. (1994, July). Generalization of Lambert's reflectance model. In *Proceedings of the 21st annual conference on computer graphics and interactive techniques* (pp. 239-246). ACM.

Oyama, T., & Yamada, W. (1978). Perceptual grouping between successively presented stimuli and its relation to visual simultaneity and masking. *Psychological research*, *40*(2), 101-112.

Paradiso, M. A., & Carney, T. (1988). Orientation discrimination as a function of stimulus eccentricity and size: Nasal/temporal retinal asymmetry. *Vision research*, *28*(8), 867-874.

Parkes, L., Lund, J., Angelucci, A., Solomon, J. A., & Morgan, M. (2001). Compulsory averaging of crowded orientation signals in human vision. *Nature neuroscience*, *4*(7), 739-744.

Pearson, J., & Clifford, C. W. (2005). Suppressed patterns alter vision during binocular rivalry. *Current Biology*, *15*(23), 2142-2148.

Peelen, M. V., & Downing, P. E. (2007). The neural basis of visual body perception. *Nature Reviews Neuroscience*, *8*(8), 636-648.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial vision*, *10*(4), 437-442.

Pirozzolo, F. J. (1979). *The neuropsychology of developmental reading disorders*. Praeger special studies.

Pirozzolo, F. J., & Rayner, K. (1980). Handedness, hemispheric specialization and saccadic eye movement latencies. *Neuropsychologia*, *18*(2), 225-229.

Rasmussen, T. and B. Milner. 1977. Clinical and surgical studies of the cerebral speech apreas in man. In: Zulch, K.J., O. Creutzfeltd and G.C. Galbraith (eds.), *Cerebral localization*. NY: Sringer Verlag, pp. 238-255.

Rayner, K. (1978). Eye movements in reading and information processing. *Psychological bulletin*, *85*(3), 618.

Renninger, L. W., & Malik, J. (2004). When is scene identification just texture recognition?. *Vision research*, *44*(19), 2301-2311.

Robertson, L. C., & Ivry, R. (2000). Hemispheric Asymmetries Attention to Visual and Auditory Primitives. *Current Directions in Psychological Science*, *9*(2), 59-63.

Ross, E. D., & Monnot, M. (2008). Neurology of affective prosody and its functional–anatomic organization in right hemisphere. *Brain and language*, *104*(1), 51-74.

Roth, E. C., & Hellige, J. B. (1998). Spatial processing and hemispheric asymmetry: Contributions of the transient/magnocellular visual system. *Journal of Cognitive Neuroscience*, *10*(4), 472-484.

Rubin, E. (1958). Figure and ground. In D. C. Beardslee & M. Wertheimer (Eds. & Trans.), *Readings in perception*, 35-101. (Original work published in 1915)

Rubin, N. (2003). Binocular rivalry and perceptual multi-stability. *Trends in Neurosciences*, 26, 289–291.

Sakano, Y., & Ando, H. (2010). Effects of head motion and stereo viewing on perceived glossiness. *Journal of Vision*, *10*(9), 15.

Sakuraba, S., Sakai, S., Yamanaka, M., Yokosawa, K., & Hirayama, K. (2012). Does the human dorsal stream really process a category for tools?. *The Journal of Neuroscience*, *32*(11), 3949-3953.

Sharan, L., Rosenholtz, R., & Adelson, E. H. (2008). Eye movements for shape and material perception. *Journal of Vision*, *8*(6), 219-219.

Sharan, L., Rosenholtz, R., & Adelson, E. (2009). Material perception: What can you see in a brief glance?. *Journal of Vision*, *9*(8), 784-784.

Sharan, L., Rosenholtz, R., & Adelson, E. H. (2014). Accuracy and speed of material categorization in real-world images. *Journal of vision*, *14*(9), 12.

Spitz, H. H., & Lipman, R. S. (1962). Some factors affecting Necker cube reversal rate. *Perceptual and Motor Skills*, *15*(3), 611-625.

Tani, Y., Araki, K., Nagai, T., Koida, K., Nakauchi, S., & Kitazaki, M. (2013). Enhancement of glossiness perception by retinal-image motion: Additional effect of head-yoked motion parallax. *PloS one*, *8*(1), e54549.

Tatler, B. W., Gilchrist, I. D., & Rusted, J. (2003). The time course of abstract visual representation. *PERCEPTION-LONDON-*, *32*(5), 579-592.

Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *nature*, *381*(6582), 520-522.

Todd, J. T., Norman, J. F., & Mingolla, E. (2004). Lightness constancy in the presence of specular highlights. *Psychological Science*, *15*(1), 33-39.

Toga, A. W., & Thompson, P. M. (2003). Mapping brain asymmetry. *Nature Reviews Neuroscience*, *4*(1), 37-48.

Tootell, R. B., Reppas, J. B., Kwong, K. K., Malach, R., Born, R. T., Brady, T. J., ... & Belliveau, J. W. (1995). Functional analysis of human MT and related visual cortical areas using magnetic resonance imaging. *Journal of Neuroscience*, *15*(4), 3215.

Toppino, T. C., & Long, G. M. (1987). Selective adaptation with reversible figures: Don't change that channel. *Perception & Psychophysics*, *42*(1), 37-48.

Tsuchiya, N., & Koch, C. (2005). Continuous flash suppression reduces negative afterimages. *Nature neuroscience*, *8*(8), 1096-1101.

Virsu, V., & Taskinen, H. (1975). Central inhibitory interactions in human vision. *Experimental Brain Research*, *23*(1), 65-74.

Wade, N. J., & Wenderoth, P. (1978). The influence of colour and contour rivalry on the magnitude of the tilt after-effect. *Vision research*, *18*(7), 827-835.

Wang, L., Weng, X., & He, S. (2012). Perceptual grouping without awareness: superiority of Kanizsa triangle in breaking interocular suppression. *PLoS One*, *7*(6), e40106.

Ward, G. J. (1992). Measuring and modeling anisotropic reflection. *ACM SIGGRAPH Computer Graphics*, *26*(2), 265-272.

Wendt, G., Faul, F., Ekroll, V., & Mausfeld, R. (2010). Disparity, motion, and color information improve gloss constancy performance. *Journal of vision*, *10*(9), 7.

Whittle, P. (1965). Binocular rivalry and the contrast at contours. *Quarterly Journal of Experimental Psychology*, *17*(3), 217-226.

Wijntjes, M. W., & Pont, S. C. (2010). Illusory gloss on Lambertian surfaces. *Journal of Vision*, *10*(9), 13.

Williams, M. A., Morris, A. P., McGlone, F., Abbott, D. F., & Mattingley, J. B. (2004). Amygdala responses to fearful and happy facial expressions under conditions of binocular suppression. *The Journal of Neuroscience*, *24*(12), 2898-2904.

Yang, E., Zald, D. H., & Blake, R. (2007). Fearful expressions gain preferential access to awareness during continuous flash suppression. *Emotion*, *7*(4), 882.

Zeki, S., Watson, J. D., Lueck, C. J., Friston, K. J., Kennard, C., & Frackowiak, R. S. (1991). A direct demonstration of functional specialization in human visual cortex. *The Journal of neuroscience*, *11*(3), 641-649.

Zhou, K., Mo, L., Kay, P., Kwok, V. P., Ip, T. N., & Tan, L. H. (2010). Newly trained lexical categories produce lateralized categorical perception of color. *Proceedings of the National Academy of Sciences*, *107*(22), 9974-9978.