Visual distortion with dynamic visual stimuli

（動的刺激による視知覚の歪みに関する研究）

區　國昌

Visual distortion with dynamic visual stimuli

Thesis by

AU Kwok Cheong Ricky

In partial fulfillment of the requirements

for the degree of Doctor of Philosophy

The University of Tokyo

Tokyo, Japan

2014

ABSTRACT

Humans use their sensory systems to interact with the fast-changing environments in which they live. Vision, a heavily relied upon sensory modality, has processing limits and is subject to inaccuracies and misjudgments. This thesis examines three specific classes of visual distortions that occur in dynamic situations. In the attentional repulsion and attraction effects, a cue that captures spatial attention can prospectively and retrospectively shift the perceived position of a target object in different ways. The positional shift has been suggested to be based on a dynamic shift of attention elicited by the cue on the target. Chapter 2 examines the necessity of visual awareness of a cue to produce the distortions. The cues, presented either before or after the target, were rendered invisible by backward visual masking. The invisible cues produced repulsion and attraction in the same directions as in the cue-visible condition, suggesting that the attentional process involved does not require visual awareness. Chapter 3 deals with the flash-lag effect (FLE), where a stationary flash appears to shift forward the perceived position of a constantly moving object, although both objects are physically aligned. The study examined the role of motion continuity on the resulting FLE. The effect magnitude was reduced, but not eliminated, when the color of the moving object alternated regularly or randomly between two colors. If the object color unexpectedly changed to a third color when the flash was presented, the FLE magnitude was largely eliminated. These data suggest that, without an unexpected change, rapid changes in object surface features somewhat degrades the maintenance of identity information, but the visual system can still register the existence of only one object in the rapidly changing stream. Chapter 4 examines numerosity estimation of objects presented in dynamic displays. Observers viewed two streams of movies that showed different numbers of dots and judged which of them contained more. Fewer dots were perceived in streams with same-color dots, compared to streams with different-color dots. This underestimation effect for same-color objects was evident only with high-speed presentation, but not in slow-speed or static displays. Furthermore, deprivation of attention by engaging the observers in a secondary task did not influence the underestimation effect. These results suggest that object substitution masking might have occurred among the same-color elements, leading to an illusion of being less numerous. Overall, these studies provided new insights regarding information processing in dynamic situations by the human visual system.

DECLRATION

I declare that this thesis is my own work and has not been submitted in any form for another degree, diploma, or qualification at any other university or institution. A portion of this thesis was written based on materials previously published as journal articles listed below, modified with permission from the respective publishers.

Au, R. K. C., Ono, F., & Watanabe, K. (2013). Spatial distortion induced by imperceptible visual stimuli. *Consciousness and Cognition, 22,* 99–110. (© 2012 Elsevier Inc.)

Au, R. K. C., & Watanabe, K. (2013). Object motion continuity and the flash-lag effect. *Vision Research, 92,* 19–25. (© 2013 Elsevier Ltd.)

Au, R. K. C., & Watanabe, K. (2013). Numerosity underestimation with item similarity in dynamic visual display. *Journal of Vision*, *13*(8):5, 1–15, http://www.journalofvision.org/content/13/8/5, doi:10.1167/13.8.5. (© 2013 ARVO)

Signed:

AU Kwok Cheong Ricky

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my supervisor, Dr. Katsumi Watanabe, for his insightful advice and guidance during the course, especially on the formulation of novel ideas, design of experiments and writing up in the different studies I engaged in. I would also like to thank the members of the thesis examination committee, Professor Michitaka Hirose, Professor Kenryu Nakamura, and Professor Yoichi Sato, for their valuable time in providing useful comments to improve this thesis. Special thanks go to Dr. Fuminori Ono for his valuable advice and support in the programming of experimental paradigm and analysis of data during the initial stage of the course. I am also grateful to Dr. Kohske Takahashi, Dr. Daisuke Matsuyoshi, and Dr. Hiroyuki Tsubomi for their helpful comments on my studies. I would also like to thank all the people in the laboratory whom I enjoyed the time with. Finally, I am thankful to my family for the unconditioned support which enabled me to complete my studies here at Tokyo.

CONTENTS

1. CHAPTER 1: INTRODUCTION

2. CHAPTER 2: SPATIAL DISTORTION INDUCED BY IMPERCEPTIBLE STIMULI

3. CHAPTER 3: OBJECT MOTION CONTINUITY AND THE FLASH-LAG EFFECT

4. CHAPTER 4: NUMEROSITY UNDERESTIMATION WITH ITEM SIMILARITY IN DYNAMIC VISUAL DISPLAY

5. CHAPTER 5: CONCLUSION

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1. General introduction

The physical environment with which humans interact is constantly changing. Humans rely on their perceptual systems to locate objects in space and to process events in their surroundings, including the visual, auditory, and somatosensory systems. The enormous amount of information in the ever-changing environment poses a significant challenge to the perceptual systems which have limited processing capacities. Given such a constraint, the perceptual systems must selectively choose information relevant to the current task or relevant to survival and neglect unimportant details, in order to construct an accurate interpretation of the environment in a limited time frame and react promptly. Resultantly, the brain obtains a rough representation of the environment through the perceptual systems, and actively constructs the remaining details to generate the most probable and rational solution (this tendency is manifested by the "Gestalt principles" in perceptual organization; Palmar, 1999; Wertheimer, 1923). The resulting percept constructed by the brain does not, then, accurately represent what is actually in the environment. In other words, our perceptual system is intrinsically bound to perceptual distortions.

Phenomena of perceptual distortions in the sensory modalities have been widely documented. Examples include illusions in auditory perception (Deutsch, 2009), in time perception (Eagleman, 2008), and even in cross-modal perception (e.g. MacDonald & McGurk, 1978; Shams, Kamitani, & Shimojo, 2000, 2002). Among the different modalities, humans rely the most on vision to interact promptly and efficiently with the world. In this thesis, I focus the discussion on visual distortions, with special emphasis on those created by stimuli presented in dynamic situations.

1.2. Visual distortions

As humans survive in the world, they need to continuously acquire visual information from the environment to guide their actions and behaviors (Müsseler, van der Heijden, & Kerzel, 2004). Examples include navigating in the correct direction toward a destination, estimating the speed and distance of surrounding objects to avoid collision, and many others. The visual system acts as an effective interface between the external world (the environment) and the internal world (the mind). However, due to the limited processing capacity of the brain, the percept created by the visual system is subject to distortions and does not accurately represent the physical reality.

In the visual system, light from the external environment enters the eyes through the lens and falls on the retina. The retina is composed of a thin sheet of photosensitive cells tuned to different frequencies of light and is situated at the back of the eyes. The physically limited number of photoreceptors on the retina poses a limit on the resolution of the resulting image. Additionally, the physically limited number of neurons in the visual processing areas of the brain also poses a challenge—the representation created and stored in the brain can never represent the full details of the physical world. Consequently, the brain has to solve an "inverse problem" (Marr, 1982; Palmar, 1999; Pizlo, 2001) of reconstructing the detailed external world, based on the inadequate information received from the image (which has low spatiotemporal resolution) projected on the retina, by relying on some "top-down" prior knowledge or assumptions.

In this chapter, I introduce three specific classes of visual distortions, including the attentional repulsion and attraction effects, the flash-lag effect, and object substitution masking, which occur in dynamic presentations. The following sections provide an overview of concepts and information relevant to the studies covered in the subsequent chapters.

1.2.1. Attentional repulsion and attraction effects

The concept of attention has been considered as a filter that selectively processes a portion of the abundance of information acquired from the environment (Broadbent, 1958), although the reality of the attentional system may not be that simple (Deutsch & Deutsch, 1963; Treisman, 1964). Studies in visual attention have distinguished the top-down (goal-directed), endogenous attentional system, from the bottom-up (stimulus-driven), exogenous attentional system (Corbetta & Shulman, 2002; Posner, 1980; Yantis, 2008). The endogenous system (also referred to as the “dorsal network”) involves the bilateral areas of the frontal and parietal regions of the brain, while the exogenous system (the “ventral network”) includes regions of the right ventral frontal and temporo-parietal cortices (Corbetta & Shulman, 2002; Fox et al., 2006). In the study of spatial attention, it is widely known that directing attention to a specific location can enhance information processing and task performance over that location (manifested as faster response times, more accurate identification of targets, and better spatial resolution) than when attention is not directed (He, Cavanagh, & Intriligator, 1996; Posner, 1980; Posner & Cohen, 1984; Posner & Petersen, 1990; Yeshurun & Carrasco, 1998).

In addition to selecting and filtering incoming information, attention has also been shown to alter perception in various dimensions of visual properties (Carrasco, Ling, & Read, 2004; Fuller & Carrasco, 2006; Gobell & Carrasco, 2005; Hikosaka, Miyauchi, & Shimojo, 1993; Liu, Fuller, & Carrasco, 2006; Turatto, Viscovi, & Valsecchi, 2007). One obvious demonstration is the spatial distortion caused by engagement of visual attention. In the attentional repulsion effect, a spatial cue that exogenously captures spatial attention is able to distort the perceived position of a subsequently presented target, so that two horizontally aligned bars of a vernier appear to drift off against each other (Suzuki & Cavanagh, 1997). Observers perceive the target bars as displaced from the location of the cue stimulus. Called the “attentional repulsion effect,” Suzuki and Cavanagh (1997) suggested that this effect was due to the cost of resource allocation to the focus of attention. In a later study, Ono and Watanabe (2011) extended the temporal window to include a range of stimulus onset asynchronies (SOA; the latency between the onset of the target and the cue) of the cue before and after the presentation of the target. The direction and magnitude of the distortion effect across the SOAs were found to resemble a sinusoidal shape, such that the strongest repulsion occurs at an SOA of around 200 ms before the target onset (i.e., SOA = -200 ms), while the strongest attraction occurs at an SOA of around 200 ms after the target onset (i.e., SOA = +200 ms) (Figures 1.1 and 1.2). According to Ono and Watanabe (2011), the dynamic shift of attention between the presentation of the cue and the target causes overshoot of the perceived stimulus positions, leading to the apparent distortion effects. As we will see in Chapter 2, such dynamic shift of attention can occur without conscious visual awareness of the observer.

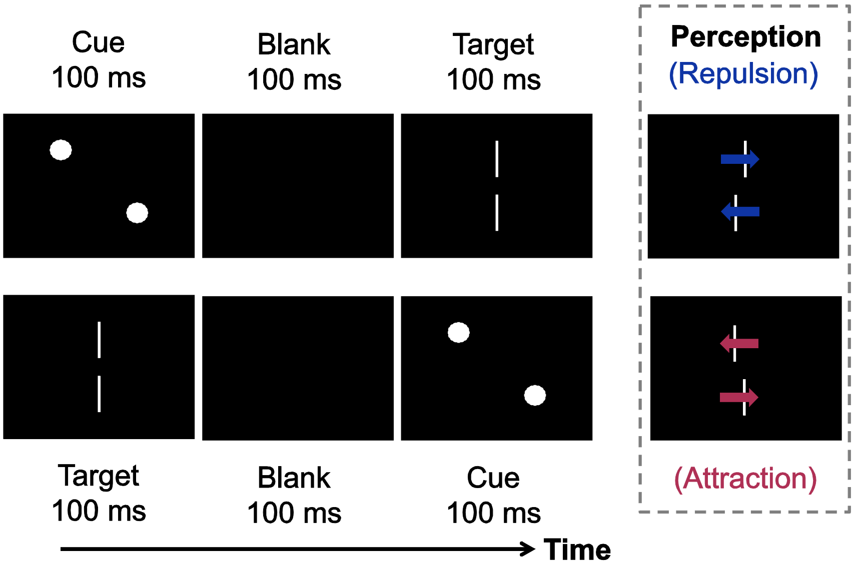

Figure 1.1. The attentional repulsion and attraction effects. In the attentional repulsion and attraction effects, the perceived position of the target appears to drift away from the cue (repulsion) when the cue is presented before the target, and appears to drift toward the cue (attraction) when the cue is presented after the target.

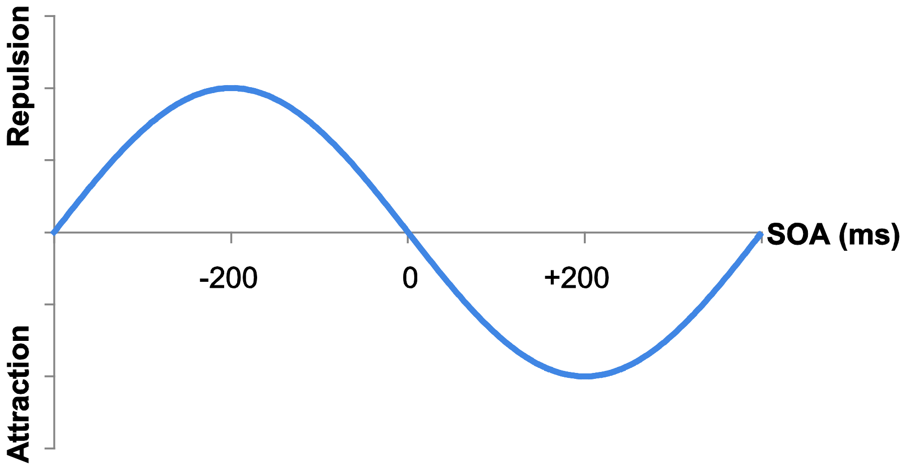

Figure 1.2. Stimulus onset asynchrony and distortion in perceived position. The attentional repulsion effect peaks when the SOA is at around 200 ms before target onset, while the attentional attraction effect is strongest when the SOA is at around 200 ms after target onset (adapted from Ono & Watanabe, 2011; figure not drawn to scale).

1.2.2. Flash-lag effect

One remarkable capability of the visual system is its ability to accurately identify the position of objects in the visual world (i.e., localization of objects). Nevertheless, psychophysical evidence has shown that under some circumstances, the visual system fails to localize objects accurately. The attentional repulsion and attraction effects introduced in the previous section are examples of failures in localization. The inaccuracy in localization is particularly noticeable when an object is moving, as evidenced by the Fröhlich effect, flash-lag effect, and others (Eagleman & Sejnowski, 2007; Whitney, 2002). The flash-lag effect (FLE; MacKay, 1958; Nijhawan, 1994) is a useful tool in studying visual mislocalization, which describes the perceptual phenomenon that the position of a briefly-flashed stationary object appears to slightly lag behind another moving object, even though the two are at physically aligned positions at the time of flash occurrence (Figure 1.3).

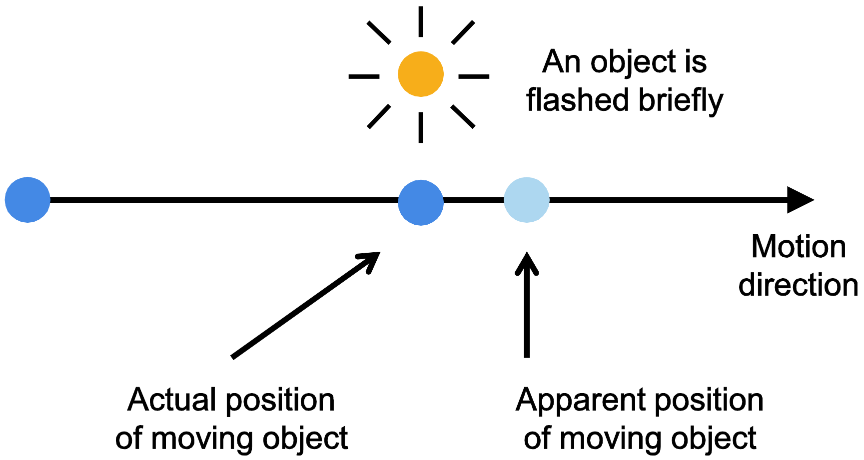

Figure 1.3. The flash-lag effect. A moving object appears to be at a position ahead of a stationary flash, although both objects are physically aligned at the moment the flash appears.

Studies on the FLE have found that this effect occurs in various conditions. For example, the FLE has been reported in objects with continuously changing features (Sheth, Nijhawan, & Shimojo, 2000), in objects moving in depth (Harris, Duke, & Kopinska, 2006; Ishii et al., 2004), in audition, and even across modalities (Alais & Burr, 2003). In addition, the FLE was found to depend on the observer's eye movements (Nijhawan, 2001) and the perceptual organization of the moving object (Watanabe, 2004; Watanabe et al., 2001).

To explain the occurrence of the FLE, a number of hypotheses have been formulated. Nijhawan (1994, 1997) proposed the motion extrapolation hypothesis, claiming that from the moment when light hits the retina, a period of time is required for processing before the object is perceived; during this period, the object has already moved to a new position, and the brain takes into account this neural delay by extrapolating the new position of the moving object. Alternatively, the latency difference hypothesis (Baldo & Klein, 1995; Kanai et al., 2009; Whitney & Murakami, 1998; Whitney, Murakami, & Cavanagh, 2000) suggests that the visual system processes moving objects more quickly than flashed objects; visual awareness is on-line and the observer can report what reaches conscious perception first. By the time the flashed object is processed, the moving object has already moved to a new position; therefore, the observer perceives the moving object as ahead of the flashed object. The third possibility was proposed by Eagleman and Sejnowski (2000). According to their idea, visual awareness is neither based on prediction of the visual system nor established on-line, but depends on the events that happen within a brief time window of around 80 ms after the flash occurs. In other words, the visual system is "postdictive" and takes into account events that occur just after a target event. Such notion of postdictive visual awareness is also consistent with the phenomenon of backward visual masking (Breitmeyer, 2007; Enns & Di Lollo, 2000), where objects presented just after the presentation of a target object can render the target object invisible to the observer. At present, the debate about the true mechanism underlying the FLE is ongoing, and it remains unclear whether any proposal can account for all the known findings related to the effect (Nijhawan, 2002). One new proposal to explain the FLE is based on the notion of object substitution masking (Enns & Di Lollo, 1997; Moore & Enns, 2004), which will be introduced in the following section. We will revisit the FLE in Chapter 3 of this thesis.

1.2.3. Visual masking and object substitution

Backward visual masking reduces the visibility of a target stimulus by the influence of an overlapping, or a nearby but non-overlapping, mask stimulus. Broadly speaking, backward visual masking can be categorized into three types: pattern masking, metacontrast masking, and the recently reported object substitution masking. In pattern masking, the mask stimulus is spatially superimposed on the position of the target stimulus. Due to the limited temporal resolution of the visual system, the mask pattern is either integrated into the target pattern (“integration masking”) or interrupts the processing of the target pattern (“interruption masking”) before the target is sufficiently processed to enter conscious perception (Bachmann & Allik, 1976; Scheerer, 1973). In metacontrast masking, the mask pattern does not overlap with the target stimulus, but just fits closely to the contours of the target without touching them (Alpern, 1953). A strong masking effect occurs when the SOA is about +50 to +100 ms (Enns & Di Lollo, 1997; Ogmen, Breitmeyer, & Melvin, 2003). In addition, physical properties of the stimulus such as brightness and contour visibility have been found to influence the strength of masking as well (Breitmeyer et al., 2006). Metacontrast masking is thought to be based on a “two-channel” mechanism. Under this framework, the onset of a stimulus initiates neural activities in a fast but short-lived channel and a slow but long-lasting channel (Weisstein, Ozog, & Szoc, 1975); masking occurs when the fast activity of the mask pattern (which is presented later) inhibits the long-lasting activity generated by the target (Bridgeman, 1971).

Enns and Di Lollo (1997) reported a new form of visual masking, named “object substitution masking,” which bears several remarkable differences from metacontrast masking. In their original experiment, the visibility of an object was impaired with the simultaneous onset of four dots surrounding it, which remained for a short time after the disappearance of the target. They proposed that the target object becomes invisible because the four-dot stimulus acts as a mask and replaces the representation of the target object. Whether the representation of the target object can reach conscious perception depends on the number of iterations of processing required to adequately process it (Di Lollo, Enns, & Rensink, 2000). The timing of the target offset is important: if the four-dot mask disappears at the same time as the target, the masking effect is eliminated. In addition, disengagement of focal attention is also an important factor: masking does not occur when the observer knows in advance the position of the target object among an array of objects (however, if the observer engages in a cognitively demanding task that involves high-level brain regions, masking would still occur even if it is spatially attended; see Dux et al., 2010); as the set size of objects in the display increases, the masking effect also increases (Di Lollo et al., 2000). This is different from metacontrast masking, as metacontrast masking remains highly effective when the target is at the center of gaze, regardless of the set size (Enns & Di Lollo, 1997). In addition, while metacontrast masking depends critically on the distance between the target and the surrounding mask pattern (the masking effect significantly decreases when separation increases; Alpern, 1953; Growney & Weisstein, 1972), object substitution masking is relatively insensitive to it. As long as the target and the mask are seen occupying the same region of the visual field, the masking effect remains robust (Di Lollo et al., 2000; Enns & Di Lollo, 1997).

Object substitution emerges when new incoming information is fed into the visual input rapidly before the system has fully processed the previous information. As described in the three-layer computational model of object substitution (Di Lollo et al., 2000; see also Enns & Di Lollo, 2000), a new visual event activates the input layer, the working space layer, and the pattern layer of the system. Reentrant connections (Bullier, McCourt, & Henry, 1988; Felleman & Van Essen, 1991) from higher extrastriate areas (of the pattern layer) allow feedback communications with the primary visual areas involved in the input and working space layers. Information stored in the pattern layer is constantly copied to the working space, and the system keeps comparing the information in the working space with that in the input layer iteratively. When there is a match of information, the signal in the pattern layer is strengthened by the weighted output from the input and working space layers, giving rise to a stable conscious perception. However, if the information at the input layer is changing rapidly, such that the pattern layer is still processing the previous information, a mismatch between the input layer and the working space will result. Such a mismatch between reentrant activities and the ongoing input causes the signal of the previous information (of which the processing is unfinished) to fade out and be replaced by the new incoming information. Whether the old information can ultimately reach conscious perception depends on the number of iterations required to complete the processing before the information has totally decayed. Therefore, if new incoming information appears rapidly enough, old information will fade out without being consciously perceived.

Such a substitution effect of objects commonly occurs when the processing system has to perform rapid updating of object information on rapidly-changing visual input (e.g., Moore & Enns, 2004). We will discuss object substitution masking in relation to two visual distortions (FLE and underestimation of object numerosity) in Chapters 3 and 4 of this thesis.

1.3. Outline of the thesis

In this thesis, I employ several effects of visual distortion to investigate how dynamic presentation of visual stimuli affects conscious perception, which gives insights into how the human visual system processes and integrates visual information in a brief period of time.

The three main themes covered in this thesis examine human visual processing in different scenarios of dynamic situations. The attentional repulsion and attraction effects (Chapter 2) concern the effect of a cue presented at a static position that affects the target at another temporal position. The FLE (Chapter 3) examines the effect of a stimulus, which is presented for a brief instant, on a target that dynamically moves with a constant velocity. The numerosity underestimation effect (Chapter 4) examines the influence of objects presented dynamically at different spatiotemporal positions on the other objects. Figure 1.4 provides a pictorial overview of them.

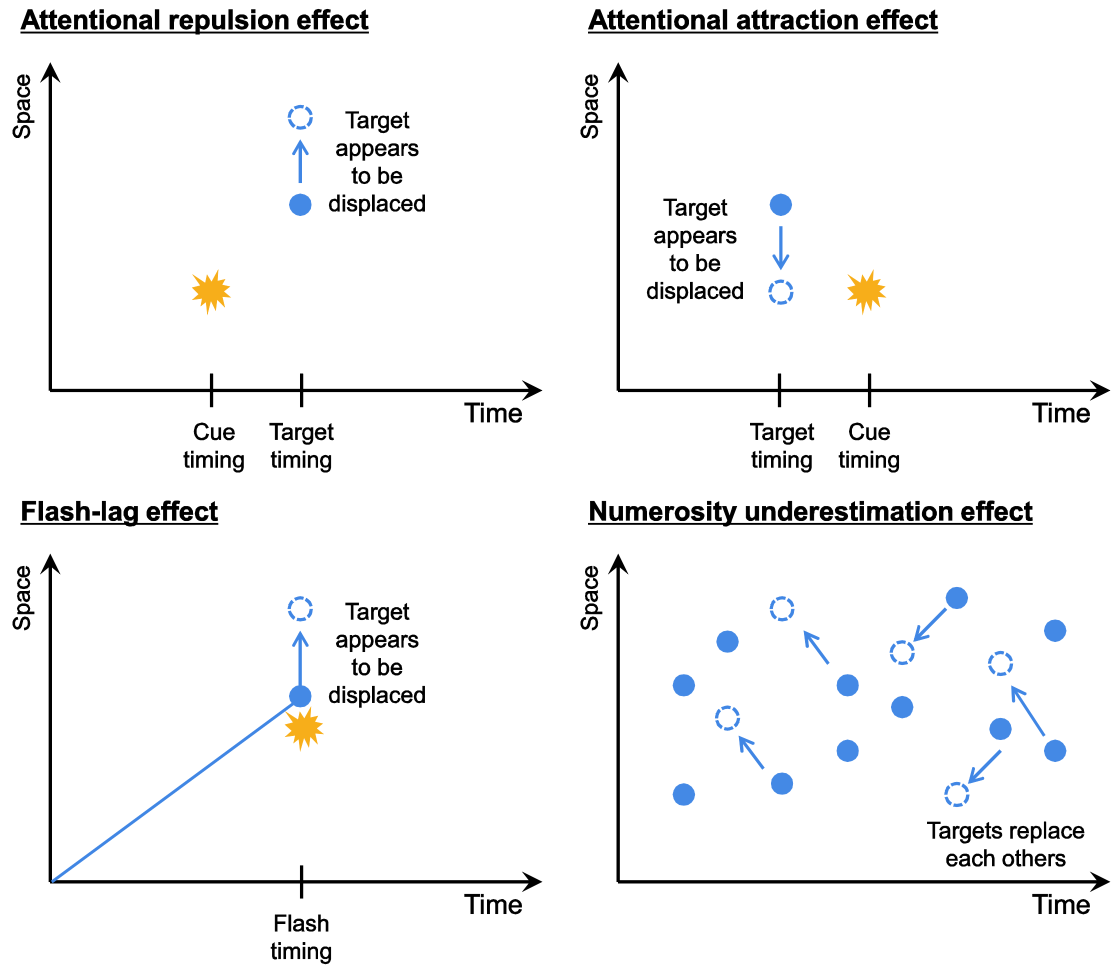

Figure 1.4. A pictorial overview of the visual distortions covered in this thesis

In Chapter 2, I study the attentional repulsion and attraction effects induced by invisible cues. The experiments showed that a spatial cue that exogenously attracts attention can alter the perceived position of a target, and this mislocalization effect still existed even when the cue was rendered invisible to the observer. In addition, under the invisible cue condition, this effect depended on the physical position, but not the observer-reported position, of the cue, thus suggesting that the mislocalization effect is based on attentional mechanisms operating below visual awareness. The study demonstrated that the visual system can process information in a postdictive manner, accounting for events occurring within a narrow time window after the target event; and the attentional processing involved can operate below conscious visual awareness.

In Chapter 3, I examine the factors that modulate the magnitude of the FLE. The FLE has been reported to be eliminated by a sudden unexpected change in the feature of the moving object that occurs during its journey. By continuously alternating the color of the moving object as it moved, the FLE magnitude was attenuated, although not totally eliminated. Further control experiments showed that the reduced effect magnitude in the changing-color condition is related to weaker perceived motion smoothness (relative to the condition without color change) and weaker perceived salience of the moving object at the time of the flash appearance (relative to the condition with the color changed only for one instant during the motion). The study showed that the visual system takes into account the history of events in dynamic presentation of a visual sequence, and the resulting perception depends on the predictability of the sequence.

In Chapter 4, I report a new visual illusion of numerosity underestimation in a dynamic presentation of objects. Objects of the same color are perceived as being less numerous than objects in the same configuration, but of different colors, when presented sequentially in a movie stream. This underestimation effect was not found when a slow presentation rate or when a static display was used. Exploitation of attentional resources affected the precision in the judgment task, but did not affect the underestimation effect in high-speed dynamic displays. This study suggests that object substitution occurs with objects in close spatiotemporal proximity. Within a brief time window, information that is inadequately processed by the visual system quickly decays and is replaced by new incoming information without reaching conscious perception.

In Chapter 5, I conclude with a summary of findings from the present studies and implications regarding the processing of dynamic visual information by the human visual system.

CHAPTER 2

SPATIAL DISTORTION INDUCED BY IMPERCEPTIBLE STIMULI

In this chapter, I examine a class of spatial distortion effect induced by exogenous capture of attention, namely the attentional repulsion and attraction effects. With behavioral experiments that employ the technique of backward visual masking, I demonstrate that the attentional shift involved in the effects does not necessitate visual awareness. The experiments are followed by discussion on possible mechanisms underlying the observations, which provides insights to the processing of information below awareness by the visual system.

2.1. Background and objectives

As mentioned in Chapter 1, spatial attention can be captured by presenting a salient stimulus (a cue) in the space to attract observer's attention. However, attentional shift is not only possible with conscious awareness of cues. In parallel with consciously perceived stimuli, the dynamic deployment of visual attention has been reported for visual stimuli that are masked and consequently go unnoticed (Mulckhuyse, Talsma, & Theeuwes, 2007). Attentional shift without visual awareness has also been employed to study the line-motion illusion (Blanco & Soto, 2009) and object-based attention (Chou & Yeh, 2011). In addition, attentional capture can also influence the events that occur after a target event. Previously, it has been demonstrated that events that occur after the offset of a target event can influence the interpretation in the visual system (Eagleman & Sejnowski, 2000). For example, the direction of the line-motion illusion (Hikosaka et al., 1993) and the attentional repulsion effect (Suzuki & Cavanagh, 1997) can be reversed by presenting the attention-capturing cue after the offset of the target

(Eagleman & Sejnowski, 2003; Ono & Watanabe, 2011), providing evidence to show that attention can exert retrospective influence on the processing of earlier visual events.

Although previous studies have shown that attentional capture can occur without awareness (e.g. Blanco & Soto, 2009; Chou & Yeh, 2011; Mulckhuyse et al., 2007), the effects of attentional capture were measured mostly by detection/discrimination performance or by reaction times. It is also still unclear whether spatial distortion can be induced via attentional shift using invisible stimuli in a similar way as that with a visible cue. In the present study, I used the attentional repulsion and attraction effects to examine the role of visual awareness of attentional cues in spatial distortion caused by the task-irrelevant visual stimuli. Here, the aim was to answer two specific research questions: first, could visual stimuli that are invisible to the observer produce spatial distortion in the same direction as clearly visible stimuli in the attentional repulsion effect? Second, could retrospective influence of attention occur without visual awareness? Specifically, would the attentional attraction effect also be observed with invisible cues? More importantly, would effects of similar magnitudes be observed under conditions with and without visual awareness? If the effect magnitude obtained under the condition without awareness is significantly attenuated, it might imply that the method of rendering the cue unavailable to conscious awareness leads to significant decrease in the strength of attentional capture, and thus the attentional repulsion and attraction effects are likely to depend largely on visual attention. Any remaining effects observed under the condition without conscious awareness would indicate that residual attention could be captured even without conscious awareness.

In Experiment 1, by presenting a brief positional cue followed instantly by a mask that made the cue invisible to the observer, the question whether the attentional repulsion and attraction effects can be reproduced with invisible cues was examined.

After obtaining affirmative results, in Experiment 2, a dual-task paradigm was employed to further investigate whether the attentional repulsion and attraction effects produced by invisible stimuli would depend on the actual position of the cue, or the position where the observer reported the cue was presented. By showing that the effects depend on the actual position of the cue (which was masked and not visible to the observer) instead of the observer-reported position (where the observer may consciously engage his/her attention at), it would provide further support that the repulsion and attraction effects obtained in the masked condition are mediated by unconscious processes.

Concerning the definitions, it would always be difficult to clearly delineate the concepts of attention and visual awareness and provide comprehensive definitions for each term, given their broadness. In the present study, visual awareness is described as the situation where the observer is able to consciously report; and attention is exogenously captured by the sudden appearance of the positional cue, which leads to differential processing engaged by the visual system on the “attended” (facilitated) and “unattended” (non-facilitated) positions (some examples were explained in the previous work by Posner and colleagues as mentioned in Chapter 1). Attention and visual awareness can obviously interact with each other, but they do not always co-exist (Koch & Tsuchiya, 2007; Lamme, 2003).

2.2. Experiment 1: attentional repulsion and attraction effects with invisible cues

2.2.1. Material and methods

2.2.1.1. Observers

Nineteen paid volunteers and the author participated in the experiment. All observers except the author were naïve as to the purpose of the study. Informed consent was obtained from the observers prior to the experiment. All of them had normal or corrected-to-normal vision.

2.2.1.2. Stimuli

Stimuli used in the present study were largely similar to that used in the previous study by Ono and Watanabe (2011). Experimental stimuli were programmed in MATLAB using the Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997). Stimuli were viewed on a CRT monitor at the refresh rate of 60 Hz controlled by a personal computer running on the Windows XP operating system. Observers viewed the stimuli at a distance of 60 cm in a dark and quiet room.

All stimuli appeared white in color (luminance = 49.02 cd/m^2) against a black background (0.016 cd/m^2) on the screen. The fixation stimulus was a dot (diameter = 0.227°) appeared at the center of the screen. The two types of cue stimuli consisted of two filled circles with diameter of 0.567° placed diagonally either in top-left and bottom-right fashion (L cue) or top-right and bottom-left fashion (R cue). Each of the circles was displaced from the center of the screen both horizontally and vertically by 1.983°. The mask stimulus consisted of four circles of the same size, placed at the four positions displaced by the same distance of 1.983° horizontally and vertically from the center of the screen; so the mask stimulus always covered the four possible positions of the circles of both the L cue and the R cue. The target stimuli were two vertical bars (length = 0.567°, width = 0.0283°) presented 1.417° above and below the center of the screen. The lower bar always appeared directly under the position of the fixation stimulus (i.e., 0° horizontal displacement), whilst the upper bar was presented randomly with -0.0283°, 0°, or +0.0283° horizontal displacement above the position of the fixation stimulus. In the present study, the reason that two non-aligned target positions were used was to confirm that the observers could correctly identify a left and right target response.

2.2.1.3. Experimental design and procedure

The experiment was carried out in three parts: Masked Cue (MC) task, Non-masked Cue (NC) task, and Cue Discrimination (CD) task. In the MC and NC tasks, the cue was presented either 200 ms before (-200 ms SOA) or after (+200 ms SOA) the appearance of the target. In each trial, the observer first saw a blank screen, and initiated the presentation of stimuli by pressing the space bar on the computer keyboard. The fixation dot then appeared for 1000 ms, followed by a blank screen lasting 100 ms. There were 3 cueing conditions for the MC and NC tasks: L cue, R cue, and no-cue control (Control cue) conditions. In the MC task, two of the cueing conditions involved a brief presentation of either the L cue or R cue for 16.67 ms (i.e., 1 frame), followed by the mask stimulus for 83.33 ms (i.e., 5 frames), resulting in a 100-ms presentation. The third cue condition involved presentation of the mask for 100 ms (i.e., 6 frames) as the no-cue control condition. Each of the three cueing conditions was equally likely to occur in each trial. The three cueing conditions in the NC task were basically the same, except that the L cue and R cue were made visible to the observer by presenting for 100 ms without mask presentation. The no-cue control condition remained the same. After the presentation of the 100-ms blank screen, in both MC and NC tasks, depending on the SOA condition of the trial, either the cue or the target was then presented for 100 ms, followed by a blank screen (100 ms), and the other stimulus (target or cue) for 100 ms. As the consequence, the SOA was controlled at either -200 ms or +200 ms in each trial. A blank screen was presented afterward, of which the observer was instructed to give a response. Observers judged whether the horizontal position of the upper bar appeared to be located to the left or to the right of the lower bar, by pressing either the Left or Right key on the keyboard. The task did not emphasize on the speed of response. Figure 2.1 depicts the flow of a trial.

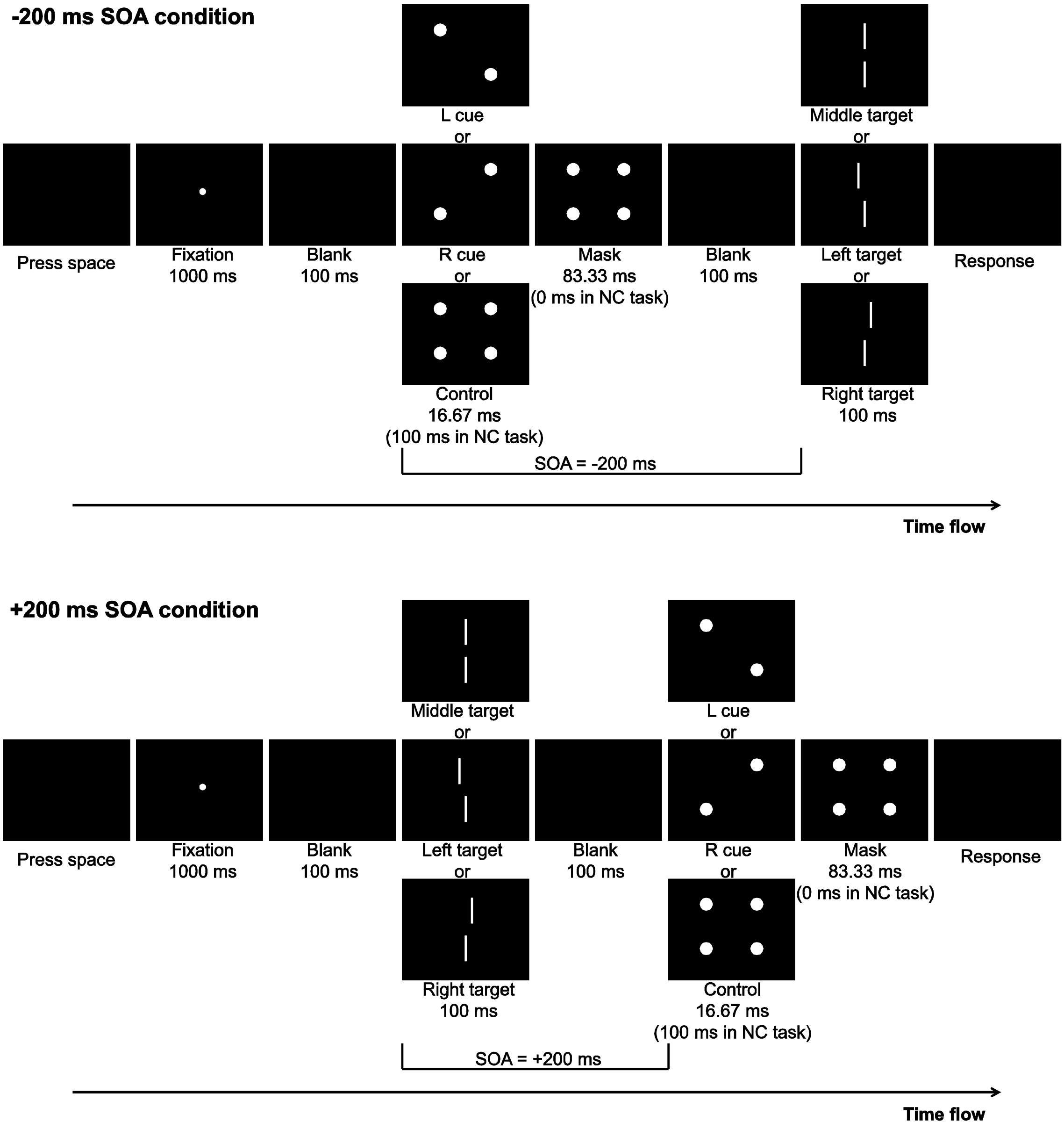

Figure. 2.1. The flow of a trial in the MC task with -200 ms (upper panel) and +200 ms (lower panel) SOA in Experiment 1. Reproduced from Au, Ono, and Watanabe (2013a) with permission.

The CD task utilized the same computer program as for the MC task, but instead of judging the position of the target, observers were asked to judge whether they saw the L cue or the R cue in each trial by pressing the Left key for seeing the L cue and the Right key for seeing the R cue. The CD task was included in order to find out whether the observers could detect the L cue and the R cue in the display used in the MC task.

Observers performed either the MC task or the NC task first, in counterbalanced order, followed by the CD task. Each of the tasks composed of trials of 2 SOA conditions × 3 cueing conditions × 3 target positions presented in pseudorandom order with 11 repetitions, resulting in a total of 198 trials for each task. The whole experiment took about 30 minutes to complete.

2.2.2. Results

2.2.2.1. Attentional repulsion and attraction with masked and non-masked cues

Figure 2.2 presents the proportion of “Right” response in each SOA condition for the NC and MC tasks (collapsed across three target conditions).

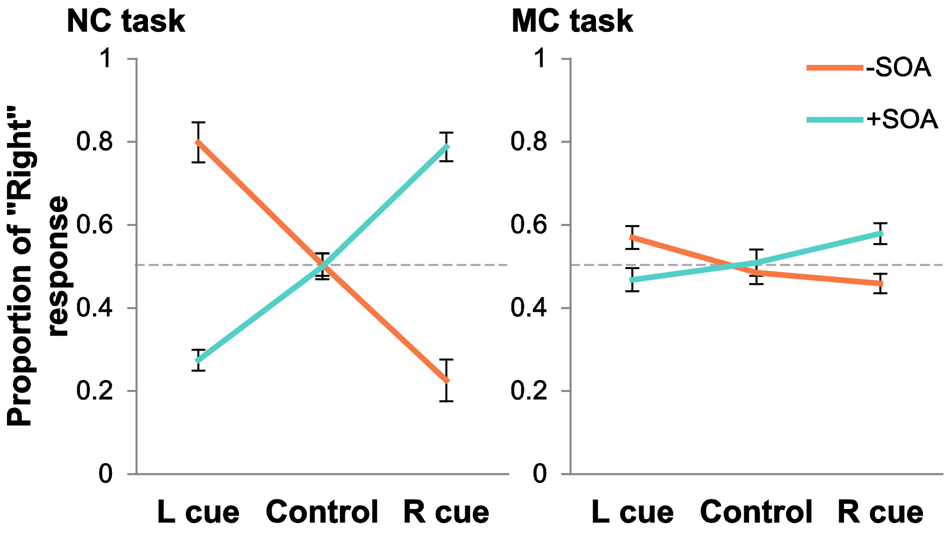

Figure 2.2. Effects of cue positions in negative and positive SOA conditions in the Non-masked Cue task (left) and the Masked Cue task (right) in Experiment 1, collapsed across three target positions; error bars represent the standard error of the mean. Reproduced from Au et al. (2013a) with permission.

A 2 × 2 × 3 × 3 (Mask × SOA × Cue × Target) repeated measures analysis of variance (ANOVA) was conducted to examine the strength of repulsion and attraction effects obtained in the NC and MC tasks. The main effects of Mask [$F(1,19) = 0.027$, $p = 0.870$], SOA [$F(1,19) = 0.518$, $p = 0.481$] and Cue [$F(2,38) = 0.686$, $p = 0.501$], and also the Mask × SOA × Cue × Target interaction [$F(4,76) = 0.592$, $p = 0.669$] were not significant; the Target main effect [$F(2,38) = 103.687$, $p < 0.001$], the Mask × Target interaction [$F(2,38) = 16.662$, $p < 0.001$], and the Mask × SOA × Cue interaction [$F(2,38) = 65.502$, $p < 0.001$] were highly significant. Analysis for the simple main effect of Target revealed significant differences across three target positions in both the NC [$F(2,76) = 119.110$, $p < 0.001$] and MC [$F(2,76) = 58.021$, $p < 0.001$] tasks. Post-hoc pairwise comparisons (alpha levels adjusted for multiple comparisons with Bonferroni correction) reported significant differences between all the pairs of target positions (i.e., L Target vs. Aligned, R Target vs. Aligned, and L Target vs. R. Target) in both the NC and MC tasks (all at $p < 0.001$). This indicates that there was no obvious asymmetric effect for the L and R target positions. The analysis of interest was the significant Mask × SOA × Cue interaction, which indicates the strength of repulsion or attraction effects concerned. Analysis on the simple interaction of Mask × Cue identified significant Mask × Cue interaction in both the -SOA [$F(2,76) = 30.730$, $p < 0.001$] and +SOA [$F(2,76) = 23.357$, $p < 0.001$] conditions. This indicates that the repulsion effect magnitude observed in the -SOA condition was significantly stronger in the NC than the MC task; attraction effect observed in the +SOA condition was also significantly stronger in the NC than the MC task. Analyses on the simple main effects of Cue revealed significant difference between the three cueing conditions in both the -SOA [$F(2,152) = 87.919$, $p < 0.001$] and +SOA [$F(2,152) = 71.041$, $p < 0.001$] conditions in the NC task, and the -SOA [$F(2,152) = 3.590$, $p = 0.030$] and +SOA [$F(2,152) = 3.352$, $p = 0.038$] conditions in the MC task. Post-hoc pairwise comparisons were conducted to examine the differences among the three cueing conditions (alpha levels adjusted for multiple comparisons). For the NC task, in the SOA condition (L cue: $M = 0.798$, $SD = 0.215$; Control cue: $M = 0.505$, $SD = 0.122$; R cue: $M = 0.226$, $SD = 0.222$), significant differences were found between all of the following pairs: L cue vs. Control cue, R cue vs. Control cue, and L cue vs. R cue (all at $p < 0.001$). The same was true in the +SOA condition (L cue: $M = 0.274$, $SD = 0.114$; Control cue: $M = 0.5$, $SD = 0.139$; R cue: $M = 0.788$, $SD = 0.152$). For the MC task, in the -SOA condition (L cue: $M = 0.570$, $SD = 0.125$; Control cue: $M = 0.485$, $SD = 0.121$; R cue: $M = 0.459$, $SD = 0.103$), significant differences were found only between the L cue vs. R cue pair ($p = 0.011$), while the differences among the L cue vs. Control cue, and R cue vs. Control cue pairs were weak ($p = 0.051$ and 0.552 respectively). Similar results were found in the +SOA condition (L cue: $M = 0.274$, $SD = 0.114$; Control cue: $M = 0.5$, $SD = 0.139$; R cue: $M = 0.788$, $SD = 0.152$): significant difference between the L cue vs. R cue pair ($p = 0.011$), and weak differences among the L cue vs. Control cue, and R cue vs. Control cue pairs ($p = 0.345$ and 0.109 respectively). The results thus demonstrated the attentional repulsion and attraction effects in both NC and MC cue conditions, though the effect magnitudes were smaller in the MC condition.

2.2.2.2. Invisibility of the masked cues

In order to ensure that the observers were unable to see the briefly presented cue during the MC task, detectability index *d’* and criterion value *c* were calculated based on the signal detection theory (Green & Swets, 1966) by computing the proportion of responses reporting R cue under the L cue condition versus the proportion of responses reporting R cue under the R cue condition in the CD task. A response was regarded as a Hit when the observer reported R cue when presented with R cue [i.e., the proportion $P(\text{Hit}) = P(\text{R}|\text{R})$] and regarded as a False Alarm (FA) when the observer reported R cue when presented with L cue [i.e., $P(\text{FA}) = P(\text{R}|\text{L})$]. The *d’* value was computed by computing the difference between the inverse of the cumulative normal distribution for $P(\text{Hit})$ and that for $P(\text{FA})$, i.e., $d' = Z[P(\text{Hit})] - Z[P(\text{FA})]$; for the criterion value, $c = \{Z[P(\text{Hit})] + Z[P(\text{FA})]\}/2$.

A one-sample *t*-test showed that the mean *d’* and *c* values of the data were both not significantly different from zero [for *d’*: $M = 0.053$, $SD = 0.345$, $t(19) = 0.690$, $p = 0.498$; for *c*: $M = 0.001$, $SD = 0.231$, $t(19) = 0.024$, $p = 0.981$], indicating that the observers were not able to reliably detect the briefly presented cues, and with no particular response bias in choosing L or R responses. Analysis on the no-cue control trials also indicated the observers did not have bias choosing L or R responses [proportion of choosing R: $M = 0.526$, $SD = 0.112$, $t(19) = 0.987$, $p = 0.337$]. The analyses therefore pointed to the idea that the attentional repulsion and attraction effects observed in the MC task were not likely to be due to residual awareness of the cues[1].

2.3. Experiment 2: do the attentional repulsion and attraction effects depend totally on the physical position of the invisible cue?

The results of Experiment 1 showed that the invisible cues were able to distort the spatial representation both prospectively and retrospectively. Then, Experiment 2 sought to further examine the involvement of attentional processes, rather than sensory processes, in the attentional repulsion and attraction effects. In Experiment 2, using the same setup as in Experiment 1, observers were required to judge both the bar position (as in the MC task) and the cue position (as in the CD task) in each trial. If the repulsion and attraction effects are based predominantly on attentional mechanisms that the invisible cue unconsciously deploys observer’s spatial attention, then the effects would only be associated with the actual position of the cue, but not with the position where the observer reports the cue exists. Conversely, if the effects are associated with the reported position of the cue, it might suggest that attentional resources are engaged at the position according to the observer’s guess, rather than the actual physical position the stimulus is presented.

---

[1] A control experiment which asked for “whether the Control Cue is presented or not” was conducted with one extra observer who indicated the Control Cue was seen in all the trials. This suggests that the observers in the Cue Discrimination task could not detect any difference in contrast energy between the L cue and R cue trials, thus further supports that the cues were totally invisible to them.

2.3.1. Material and methods

2.3.1.1. Observers

The author and nineteen observers who were naïve as to the purpose of the study participated. All observers had normal or corrected-to-normal vision. Informed consent was obtained before the experiment. Except for the author, no observers had participated in Experiment 1.

2.3.1.2. Stimuli and procedure

The stimuli and the viewing condition used were the same as those in Experiment 1. However, in addition to responding to the target bar location, the observers were required to give response regarding the cue location (L cue or R cue) after giving response on the bar position in each trial. They were asked to press the Left or Right key regarding the bar position first, and then the Up or Down key (for L cue and R cue respectively) for the cue position. As in the MC task of Experiment 1, a total of 198 trials were presented (2 SOA conditions × 3 cueing conditions × 3 target positions × 11 repetitions). The experiment took about 15 minutes to complete.

2.3.2. Results

2.3.2.1. Attentional repulsion and attraction effects in the dual-task paradigm

Similar to Experiment 1, the proportion of “Right” responses in each SOA and target position condition was computed, and is plotted against the cue conditions, as shown in Figure 2.3 (collapsed across three target positions).

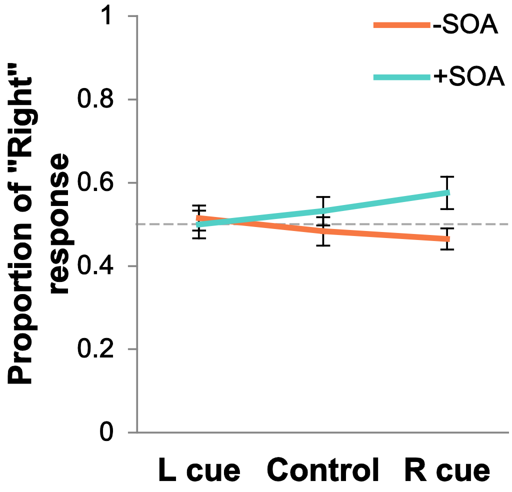

Figure 2.3. Effects of cue positions in the negative and positive SOA conditions for the dual task in Experiment 2, collapsed across three target positions; error bars represent the standard error of the mean. Reproduced from Au et al. (2013a) with permission.

A 2 × 2 × 3 (SOA × Cue × Target) repeated measures ANOVA was performed to examine the strength of repulsion and attraction effects obtained in the dual task. The analysis indicated no significant main effect of SOA [$F(1,19) = 2.532$, $p = 0.128$] or Cue [$F(2,38) = 0.554$, $p = 0.579$], but there were significant main effect of Target [$F(2,38) = 45.954$, $p < 0.001$], SOA × Target interaction [$F(2,38) = 4.434$, $p = 0.019$], and SOA × Cue interaction [$F(2,38) = 7.374$, $p = 0.002$]. The SOA × Cue × Target interaction was not significant [$F(4,76) = 0.934$, $p = 0.449$], indicating that there was no obvious effect of target position on the repulsion/attraction effect observed. For the significant Target main effect, post-hoc analysis was performed and significant differences were found between all the pairs of target positions: L Target vs. Aligned, R Target vs. Aligned, and L Target vs. R Target (all at $p < 0.001$), implying that there was no obvious asymmetric effect of target position. The analysis then focused on the significant SOA × Cue interaction to explore the strength of repulsion or attraction effects. For the simple main effects of Cue, analyses revealed significant difference between the three cueing conditions in the +SOA condition [*F*(2,76) = 6.188, *p* = 0.003], and the difference was close to significant in the -SOA condition [*F*(2,76) = 2.739, *p* = 0.071]. As in Experiment 1, post-hoc pairwise comparisons were carried out to examine the differences among the three Cue conditions. For the -SOA condition (L cue: *M* = 0.515, *SD* = 0.136; Control cue: *M* = 0.483, *SD* = 0.154; R cue: *M* = 0.465, *SD* = 0.114), the difference between the L cue vs. R cue pair was close to significant (*p* = 0.023), while the differences between the L cue vs. Control cue and R cue vs. Control cue pairs were weak and not significant (*p* = 0.145 and 0.403 respectively). Similar results were found in the +SOA condition (L cue: *M* = 0.5, *SD* = 0.148; Control cue: *M* = 0.532, *SD* = 0.153; R cue: *M* = 0.576, *SD* = 0.174), with significant difference found between the L cue vs. R cue (*p* < 0.001) pair, but not for the L cue vs. Control cue and R cue vs. Control cue pairs (*p* = 0.145 and 0.046 respectively).

Regarding the visibility of the masked cue, the proportion of L cue trials in which observers responded “Left” and the proportion of R cue trials in which observers responded “Right” were calculated and compared against the hypothesized value of 0.5 using one-sample *t*-tests. The observers were neither able to reliably report the L cue [*M* = 0.531, *SD* = 0.0861, *t*(19) = 1.614, *p* = 0.123] nor the R cue [*M* = 0.489, *SD* = 0.105, *t*(19) = -0.452, *p* = 0.656]. The mean *d’* value was also not significantly different from zero [*M* = 0.0530, *SD* = 0.286, *t*(19) = 0.827, *p* = 0.419], showing that the observers could not reliably detect the cue location. In addition, the mean criterion value *c* was not significantly different from zero [*M* = 0.0539, *SD* = 0.202, *t*(19) = 1.196, *p* = 0.247], showing that the observers did not exhibit particular bias in choosing “Left” or “Right” responses. This was also reflected in the observers’ proportion of choosing “Left” or “Right” responses in the no-cue control trials [proportion of “Right” response: *M* = 0.508, *SD* = 0.131, *t*(19) = 0.258, *p* = 0.799].

In Experiment 1, observers' visibility on the masked cue was measured in a separate block of trials (i.e., the CD task), leaving open the possibility that some "residual" visibility might be in effect when they performed the bar judgment task (i.e., the MC task), and this might induce the small repulsion and attraction effects observed. To exclude this possibility, the repulsion and attraction effects (in terms of the proportion of giving "Right" response on target position) was calculated separately for trials in which the observers reported the cue positions correctly and incorrectly in Experiment 2 (Figure 2.4, collapsed across three target positions). If the cues were totally invisible to the observers, repulsion and attraction effects would be shown in the appropriate directions in the incorrectly reported trials.

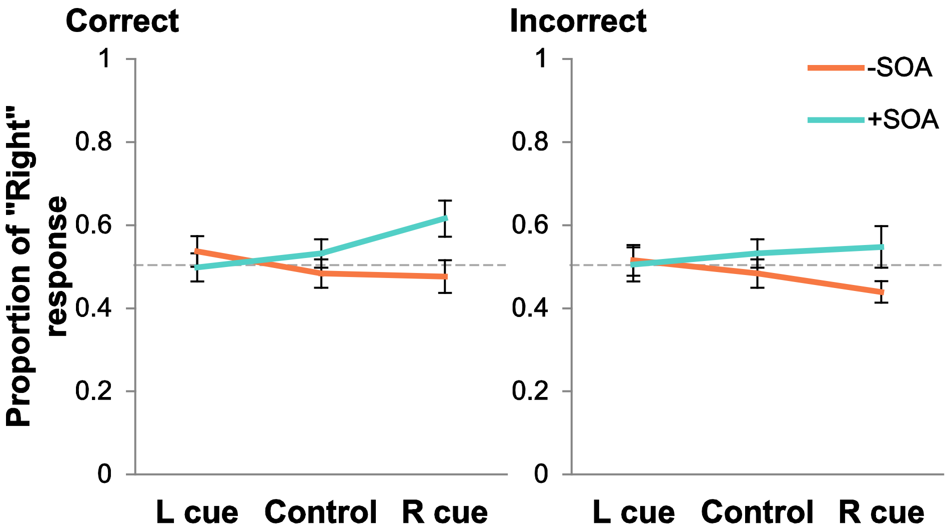

Figure 2.4. Effects of cue positions in the negative and positive SOA conditions for the dual task of Experiment 2 calculated separately for correct and incorrect trials, collapsed across three target positions; error bars represent the standard error of the mean. Reproduced from Au et al. (2013a) with permission.

The 2 × 2 × 3 (SOA × Cue × Target) repeated measures ANOVA was conducted separately for the sets of trials that the position of the cue was correctly, and incorrectly reported. Among the correctly reported trials, results reported significant main effect of Target [$F(2,38) = 26.575$, $p < 0.001$], and also significant SOA × Cue interaction [$F(2,38) = 6.405$, $p = 0.004$]. No significant SOA × Cue × Target interaction was detected [$F(4,76) = 1.506$, $p = 0.209$]. Simple main effect analysis on SOA showed significant effect of Cue in the +SOA condition [$F(2,76) = 5.427$, $p = 0.006$], but not in the -SOA condition [$F(2,76) = 1.619$, $p = 0.205$]. This indicated that the attraction effect was relatively strong in the +SOA condition, while the repulsion was too weak to be detected in the -SOA condition, among the trials with correctly reported cue location. Following the significant simple main effect of Cue in the +SOA condition, post-hoc analysis (with alpha levels adjusted by Bonferroni correction) showed significant difference between the L cue vs. R cue pair ($p = 0.002$), and a non-significant difference for the R cue vs. Control cue pair ($p = 0.049$). Similarly, for the incorrectly reported trials, significant main effect of Target [$F(2,38) = 37.857$, $p < 0.001$] as well as a significant SOA × Target interaction [$F(2,38) = 6.849$, $p = 0.003$] were found. However, the interaction of interest (the SOA × Cue interaction) was found to be not significant [$F(2,38) = 2.245$, $p = 0.120$] among the trials with the cue location incorrectly reported. Based on this result, one may conjecture that concurrently performing the target location judgment task and the cue judgment task might have weakened subjects' focus on the target judgment task, thus leading to weakened repulsion and attraction effects obtained. Nevertheless, the directions of the effects were consistent with predictions for each SOA condition.

2.3.2.2. Attentional effects associated with reported cue position

The dual-task nature of Experiment 2 made it possible to explore the strength of attentional effects according to observers' report regarding the cue location. The magnitudes of effects for each SOA condition associated with the reported positions (L cue or R cue), instead of the actual cue position (L cue, Control cue, or R cue) which was used in the previous analyses, of the masked cue for both the cued and the no-cue control conditions were computed. By analyzing the data in this way, if the repulsion and attraction effects observed in each SOA condition occur in the expected directions (i.e., repulsion for -SOA and attraction for +SOA) according to the actual physical position where the cue was at (which is invisible to the observers) rather than the reported position (where the observer believes or guesses the cue was located at), this might possibly imply that the cue captured attention to its location below conscious awareness. Figure 2.5 shows the magnitude of the effects.

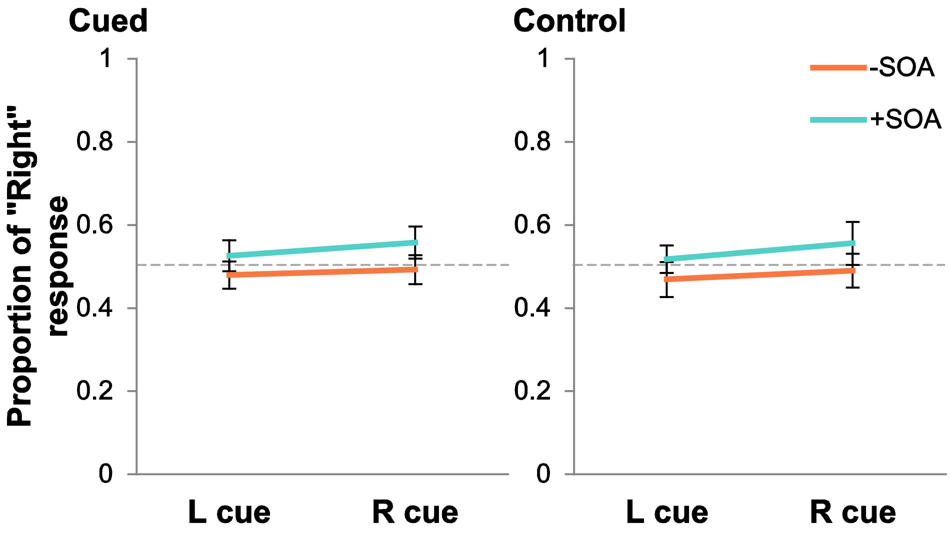

Figure 2.5. Effects of cue positions in the negative and positive SOA conditions for the dual task of Experiment 2 calculated using observer-reported cue positions, with individual magnitudes for cued and no-cue control conditions, collapsed across three target positions; error bars represent the standard error of the mean. Reproduced from Au et al. (2013a) with permission.

The 2 × 2 × 3 (SOA × Reported-Cue × Target) repeated measures ANOVA was separately performed on the data of the cued trials (i.e., with the masked L cue or R cue actually presented) and the no-cue control trials (i.e., with the Control cue presented). In the cued trials, results of the ANOVA revealed significant main effect of Target [$F(2,38) = 48.306$, $p < 0.001$]. All other main effects and interactions were not significant, including the SOA × Reported-Cue interaction of interest [$F(1,19) = 0.687$, $p = 0.418$]. This implied that the repulsion and attraction effects did not occur according to the observer-reported cue location. In the same vein, in the no-cue control trials, the Target main effect [$F(2,38) = 16.266$, $p < 0.001$] was found to be statistically significant. The SOA × Target interaction was also significant [$F(2,38) = 4.224$, $p = 0.022$]. Similar to the cued trials, the SOA × Reported-Cue interaction was not significant [$F(1,19) =$

0.2022, p = 0.658], suggesting again that the repulsion and attraction effects did not occur with the reported cue locations. Hence, these results support the idea that the attentional repulsion and attraction effects were primarily based on the physical (yet unaware) position of the cue where the observer's attention was drawn, but not where the observer subjectively believed he/she saw the cue.

2.4. Discussion

The present study examined whether the attentional repulsion and attraction effects would be observed with masked (and therefore invisible) cues. The results showed that, although the effect magnitudes were smaller, the invisible cues produced spatial distortion in the same manner as the visible cues (Experiment 1). Furthermore, the spatial distortion effects were associated with the actual presented position of the cue, instead of the position at which the observers reported the cue was presented (Experiment 2). These results suggest that attentional modulation of spatial representation was not necessarily associated with visual awareness of attention-deploying stimuli, and that the processing of attentional deployment by imperceptible visual stimuli is based on the physical existence of the stimulus, rather than on subjective guess about its location, in nature.

While several studies examining the effects of imperceptible stimuli on attentional processes have been reported, the present study is the first to show attentional effects on spatial distortion by imperceptible cues. If the attentional repulsion and attraction effects induced by visible cues in previous studies were purely due to perceptual principles such as apparent motion instead of attentional principles, no distortion effect would be observed under the condition with invisible cues. The finding that spatial distortion occurs without visual awareness of the cues supports the idea that the repulsion and attraction effects are due to attentional rather than perceptual process (Eagleman & Sejnowski, 2007; Ono & Watanabe, 2011; Pratt & Arnott, 2008; Suzuki & Cavanagh, 1997). The role of cue visibility might be simply to boost the strength of attention drawn, leading to the stronger attentional effects observed in the non-masked cue condition. The possibility also remains open that the reduced magnitude of effects observed in the masked condition was due to reduced contrast energy of the cues (Talgar, Pelli, & Carrasco, 2004; Watson, Barlow, & Robson, 1983), since the cue was presented for only 17 ms in the masked condition but for 100 ms in the unmasked condition.

2.4.1. Recurrent processing in visual awareness

To explain the present results, the model of recurrent processing of visual awareness described by Lamme (2003) may worth particular discussion.

Visual attention and visual awareness are both selective in nature and are easily confused concepts (Koch & Tsuchiya, 2007; Lamme, 2006). In a review paper explaining how visual awareness is different from attention, Lamme (2003) described a model of awareness that clearly distinguished the two processes. In the model, as visual input comes in, neurons in the early visual areas are first activated, and the feed-forward sweep of signals rapidly projects the information to the extrastriate visual areas including both the parietal and temporal cortices, to extract visual features such as shape, color, and motion (Lamme, 2003, 2006). But such processing is still at the unconscious stage. After a short latency, neurons in the activated regions begin to form local networks that perform recurrent processing. It is the formation of recurrent connections that determines whether the representation of a visual stimulus reaches the conscious level; backward masking of the stimulus could suppress such formation (Lamme, Zipser,

& Spekreijse, 2002) and thus keep the processing at the unconscious level. Empirically, it has also been shown using transcranial magnetic stimulation that recurrent feedback signals from higher visual areas such as V5 to the primary visual area (V1) are crucial for visual awareness of motion (Pascual-Leone & Walsh, 2001). As local recurrent interactions extend to widespread areas and incorporate higher-level frontal executive regions, the stimulus reaches “access awareness” (Block, 2005) and can be stored in working memory. Consequently, conscious reporting becomes possible even after the removal of the attended stimulus. Information about unattended stimuli remains restricted to local interactions in early areas. In the same vein, activations in the higher cortex such as the prefrontal and parietal areas have been shown to be important for subjective visual consciousness in human neuroimaging studies (Lau & Passingham, 2006; Rees, Kreiman, & Koch, 2002; Thompson & Schall, 1999). For top-down attention, Koch and Tsuchiya (2007) also provided examples of percepts and behaviors arguing that attention and consciousness are dissociable and need not occur together, such as iconic memory, pop-out in search (in which top-down attention is not required), priming, and adaptation (in which consciousness does not arise).

Based on the above framework, in the present study, processing of the masked cue remained at the unconscious level (i.e., failed to reach access awareness). The widespread formation of recurrent interactions between neurons might have been disrupted by masking and thus was not established (Macknik & Livingstone, 1998; see also Supèr, Spekreijse, & Lamme, 2001). As a result, the information regarding the existence and position of the cue failed to reach access awareness and the observers therefore could not report it. Nevertheless, at the behavioral level, the results demonstrated the attentional repulsion and attraction effects in the same directions whether processing took place at the conscious or unconscious levels. Hence, it is likely that the observed effects in both the NC and MC tasks shared the same attentional process; that is, focal attention devoted to attended locations over unattended locations creates differential spatial distortion patterns arising from predictive and postdictive influences. This is also in line with previous findings that subliminally presented stimuli can draw spatial attention even if they are beyond the observer’s explicit awareness (Astle, Nobre, & Scerif, 2010). Recent psychophysical findings also showed that conscious visual awareness is not necessary for a subliminal stimulus to attract observer’s attention, and the processes involved in visual awareness and attention are somewhat independent (Hsieh, Colas, & Kanwisher, 2011). With regard to attentional influence in the present study, under Lamme’s (2003) framework, the engagement of attention (no matter consciously or unconsciously engaged) on cued positions might probably cause a trace of locally activated neurons to become ready to react more efficiently and strongly to the positions where the cue attracted attention to. Such enhancement on these pre-activated neurons which are responsible for the attention-engaged receptive fields might lead to prospective (in the -SOA condition) and retrospective (in the +SOA condition) differential processing (through more efficient formation of local recurrent interactions) at these positions than the positions without attentional engagement when the target bars were presented, resulting in the apparent visual distortion on the position of the target bars. Note that the formation of local recurrent interactions induced by attentional engagement can remain at the level of “phenomenal awareness” (Block, 2005) that the observer is unable to consciously report, as long as the recurrent interaction is not widespread and reaches the higher cortical areas which leads to access awareness. Therefore, engagement of attention does not necessitate (access) awareness.

2.4.2. Dynamic attentional shift without visual awareness

The present study is also the first to demonstrate that the retrospective effect of attention can occur without awareness. The retrospective influences of attention may be explained by the dissociation between attention and awareness. If attention is deployed to a cued location without awareness of the cue, and if attentional shift is much quicker than the establishment of visual awareness (Libet et al., 1964, 1979), then the effect might appear to occur retrospectively. In the case of the attentional attraction effect, the appearance of the target is accompanied by rapid deployment of attention. The presentation of cue after a short latency triggers attentional shift. If awareness is not a necessary condition for initiating attentional processes (as the results of the present study suggest), the attentional dynamics could operate and complete at a moment before visual awareness is established (see also Au et al., 2013b). At the time when awareness of the target is established (which some researchers suggested it might require about 500 ms; Libet et al., 1964), the visual system creates a mental representation of the target. By that time, the process of attentional shift has already completed and is ready to act directly on the representation of the target, resulting in the apparent position shift. The dynamic shift of attentional focus from the target to the cue results in a shift of perceived target position in the same direction, resulting in the attentional attraction effect (Chien, Ono, & Watanabe, 2011). This idea is basically in accordance with the combined idea that visual awareness requires some time to develop after stimulus presentation, and that events occurring after the disappearance of the visual stimulus can influence the visual system’s interpretation and its processing is delayed for a brief window of time (Eagleman & Sejnowski, 2000, 2003; Ono & Watanabe, 2011), therefore attentional influence could occur before and/or without visual awareness

(Blanco & Soto, 2009; Hikosaka et al., 1993; Mulckhuyse et al., 2007; Suzuki & Cavanagh, 1997).

2.5. Conclusion

Using brief, imperceptible positional cues, this study demonstrated that visual awareness of a cue is not a necessary condition for attentional modulation of spatial representation to occur. The results showed the occurrence of attentional repulsion and attraction effects even without conscious visual awareness of the positional cues, suggesting that the effects are due to attentional rather than perceptual processing. Visual awareness is dissociable from attentional processing on its role in spatial distortion.

In this chapter, we examined the effect of a cue presented at a static position which dynamically affects the target at another time point. In the next chapter, we look into another class of mislocalization effect─the FLE, which occurs when a cue presented for a brief instant affects the target that dynamically travels at a constant velocity in space.

CHAPTER 3

OBJECT MOTION CONTINUITY AND THE FLASH-LAG EFFECT

In this chapter, I examine a visual illusion of spatial mislocalization induced by a briefly-flashed stimulus, namely the flash-lag effect (FLE). In the experiments described here, by alternating the color of a moving object while it is moving along a uniform trajectory, I investigated the role of perceived motion smoothness and salience of object feature on the magnitude of the FLE. The data are then discussed in relation to the history of change in object surface feature and the maintenance of object identity information in dynamic motion by the visual system.

3.1. Background and objectives

As reviewed in Chapter 1, the frequently discussed explanations for the FLE in the literature include motion extrapolation (Nijhawan, 1994), latency difference (Baldo & Klein, 1995; Whitney & Murakami, 1998), and postdiction (Eagleman & Sejnowski, 2000). Nevertheless, without a consensus, the ongoing debate is yet to be settled. Moore and Enns (2004) proposed a new explanation of the FLE by viewing the effect as the result of an ongoing object updating process based on the principle of object substitution (Enns & Di Lollo, 1997). They proposed that, due to the ongoing updating process, positional information of the moving object acquired immediately after the flash presentation overwrites (replaces) that acquired at the time of the flash presentation, resulting in the illusory perception that the moving object overshoots the flash. In the case where the moving object stops at the time of the flash presentation, since there is no new information about the moving object after the flash presentation that can replace the previous information, the alignment of the two objects can be perceived accurately. In the same study, Moore and Enns (2004) further reported that, when the visual features of the moving object, such as size and color, change abruptly at the moment of the flash presentation and changed back immediately after it (this is referred to as the “One Change” motion condition in the present study), observers tend to perceive the moving object appearing at two positions (one object with the changed color and aligned with the flash, and the other with the original color located at a position in front of the flash) when asked about the perception at the moment of the flash presentation. The authors explained that the disruption of motion continuity by a large and transient change leads the visual system to interpret the scene as containing two separate objects. When the original object reappears at a new position after the flash has disappeared, its position and color information is updated, while the information acquired at the moment of the flash presentation (which is interpreted as a different object) is spared from the overwriting process. However, if a scene-based reason is provided for the discontinuity, the object updating process would be spared from disruption, preserving the representation of the original object, and thus the FLE can be observed (Moore, Mordkoff, & Enns, 2007).

According to the idea above, whether object motion continuity can be preserved depends on whether only a single (i.e., the same) object is identified throughout the motion scene. The nature of object persistence has been widely studied with the approach of object file theory (Kahneman, Treisman, & Gibbs, 1992). According to this theory, episodic representations (object files) keep track of individual entities in the scene over space and time, and are updated based on spatiotemporal information (i.e., location at different moments). Object files store the representations of persistent objects and mediate conscious perception, informing the observer about “which went where” (Mitroff, Scholl, & Wynn, 2005); and such representations can be quite persistent to store object identity information on a scale of seconds (Noles, Scholl, & Mitroff, 2005). Empirical evidence has suggested that object files encode identity information rather than semantic or precise physical information (i.e., physical features) of objects, and that object file representations are flexible (Gordon & Irwin, 1996, 2000). While Mitroff and Alvarez (2007) have shown that spatiotemporal information, but not surface features, effectively determines object persistence, Moore, Stephens, and Hein (2010) demonstrated that object feature alone could determine object persistence under some conditions. It is therefore still unclear what role object surface feature plays in the establishment and maintenance of object files.

An interesting question that can be derived from the study of Moore and Enns (2004) is: what will be observed in a stream of events consisted of an object moving in a uniform trajectory while its surface feature (e.g., color) keeps changing? This would represent a case in which spatiotemporal continuity suggests only a single object is moving throughout the journey, but the information from surface feature suggests that multiple units exist. In the present study, I investigated this question by introducing two conditions—Alternating stream (in which the color of the moving object alternates between two colors; Experiment 3) and Random stream (in which the color of the moving object changes randomly between two colors; Experiment 4), in addition to the One Change and No Change conditions employed in the original study by Moore and Enns (2004). Based on previous work on the object file theory, if spatiotemporal information dominates the formation and updating of episodic object files (so that the visual system identifies only one object in the stream), we would expect the FLE to occur even in the Alternating and Random stream conditions. This would also mean that the unexpected and highly salient change at the moment of flash presentation in the One Change stream is a necessary condition for breaking motion continuity (such that the visual system identifies multiple objects in the stream) and eliminating the FLE. Alternatively, if object surface feature plays a significant role in maintaining object files, the history of color change along with the motion would cause the visual system to identify that multiple objects exist in the motion stream. In this case, the FLE might be eliminated because the overwriting process of previous information at each instant is largely disrupted by the color change.

3.2. Experiment 3: color alternations and the magnitude of flash-lag effect

To examine the effect of object motion continuity on the resulting magnitude of FLE, the performance across three motion streams of stimuli (No Change, One Change, and Alternating) were compared. In the Alternating stream, the disc color alternated regularly between red and green as the disc moved along the trajectory. Since the color change was regular, the observer could fairly predict how the disc would behave as it moved.

3.2.1. Material and methods

3.2.1.1. Observers

Twelve paid volunteers participated in the experiment. All were naïve as to the purpose of the study and had normal or corrected-to-normal vision. Informed consent was obtained prior to the experiment.

3.2.1.2. Stimuli and procedure

The stimuli used in the experiment were developed based on the study by Moore and Enns (2004; Part 2), and were programmed in MATLAB R2012b (MathWorks, USA) using the Psychophysics Toolbox extension (version 3.0.8; Brainard, 1997; Pelli, 1997). The stimuli were displayed on a CRT monitor with a refresh rate of 100 Hz (resolution = 800 × 600 pixels), controlled by a personal computer running the Windows 7 operating system. Observers viewed the stimuli at a distance of 60 cm in a dark and quiet environment.

All experimental stimuli were presented on a black background (luminance = 0.022 cd/m^2). The observer initiated each trial by pressing the space bar on the keyboard. As the space bar was pressed, a white fixation cross consisting of a horizontal line and a vertical line (length = 0.317°, width = 0.0453°) appeared at the center of the screen and remained throughout the trial until a response was given. Observers were required to remain fixated on the cross throughout the trial. At the same moment when the trial was initiated, a circular target stimulus (diameter = 0.907°) in either red or green (luminance = 0.47 cd/m^2) appeared either directly above or directly below the fixation cross at a distance of 4.171° and stayed for 500 ms. After that, the target stimulus started to move in clockwise or counter-clockwise direction on an imaginary circle (radius = 4.171°) around the fixation cross for a random angular distance of 105°, 195°, 285°, or 375° at an angular speed of 15°/frame. Each frame was shown on the screen for 70 ms, and thus the duration of the motion stream was 490 ms, 910 ms, 1330 ms, or 1750 ms, respectively. One of the following three possible motion streams was presented in each trial: No Change, One Change, or Alternating. In the No Change stream, the color of the target remained unchanged throughout the trial. In the One Change stream, the target color changed to the other color at the second last frame of the motion (which corresponded to the position directly above, below, to the left, or to the right of the fixation, and thus was always aligned with the fixation), and changed back to its original color in the last frame of the motion. In the Alternating stream, the color of the target alternated between red and green in each frame of the motion (Figure 3.1).

The flash stimulus was a white disc (diameter = 0.544°, luminance = 2.89 cd/m^2) presented at the position directly above, below, to the left, or to the right of the fixation at a distance of 2.901°. For presentation timing, the flash was presented for only one frame at the third last, the second last, or the last frame of the motion. These three flash conditions resembled the “behind,” “aligned,” and “ahead” conditions adopted in the previous study (Fig. 1a and 1b of Moore & Enns, 2004). In addition to these three flash conditions, there were also two baseline flash conditions for each stream condition. In the previous study, when the flash appeared, the target disc was presented at the second last position of the motion in the No Change condition, and was presented at the second last and the last position of the motion in the One Change condition (Fig. 1c of Moore & Enns, 2004). However, in the present study, both of these baseline conditions were included in all stream conditions to reduce any possible difference in the magnitude of the FLE or bias elicited by the different baseline conditions used in the No Change and One Change streams, thus allowing a better comparison across different stream conditions. To make the descriptions clear, in the Baseline 1 condition, the target stimulus stream was identical to the “aligned” flash condition, except that the target disc disappeared along with the flash; in the Baseline 2 condition, the target stimulus stream was essentially the same as the Baseline 1 condition, except that an additional disc was also presented in the second last frame and disappeared along with the flash. This additional disc was presented at the position where the disc should appear in the last frame in a non-baseline condition (see the “small change” and “large change” conditions in Fig. 1c of Moore & Enns, 2004). Therefore, in the two baseline conditions, the target discs were presented up to the second last frame of the motion stream, and only the central fixation cross was displayed in the last frame (Figure 3.2).

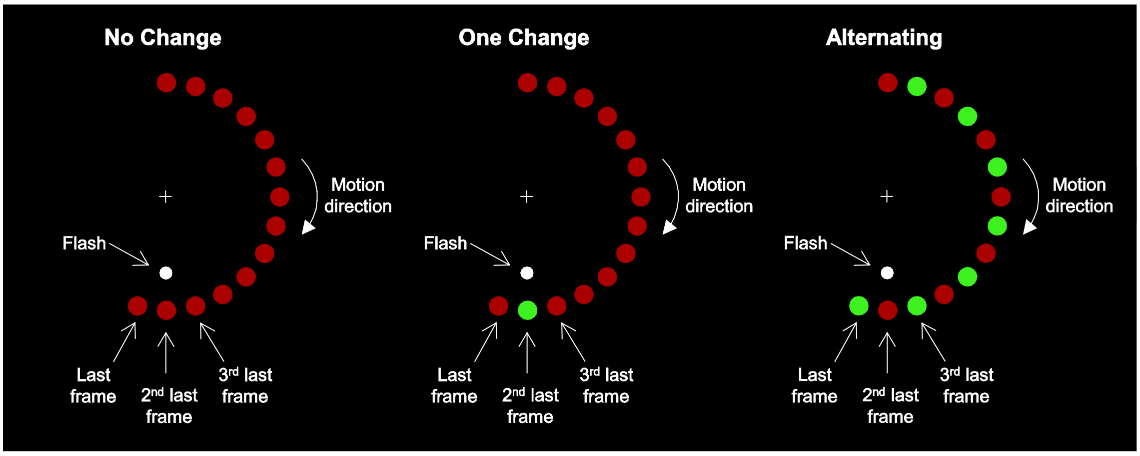

Figure 3.1. The three motion stream conditions employed in Experiment 3. Reproduced from Au and Watanabe (2013a) with permission.

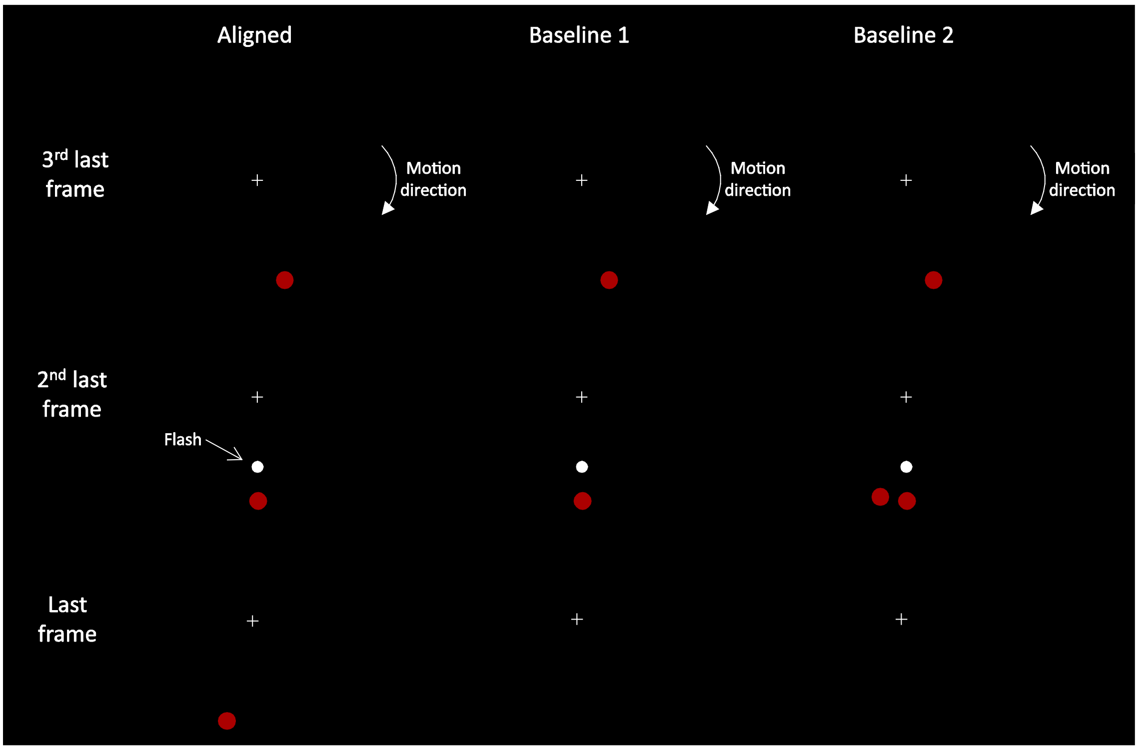

Figure 3.2. An illustration of the Baseline 1 and Baseline 2 conditions.

Observers were required to judge, upon the disappearance of the target disc, whether the target disc was in alignment with the flash (and also the fixation) at the moment when the flash appeared. Following Moore and Enns (2004), the observers were also instructed to give the response of "aligned" if they saw two target discs and either one of them was aligned with the flash. There were a total of 480 trials (3 Streams conditions × 5 Flash conditions × 2 Starting Positions × 4 Travel Distances × 2 Starting Colors × 2 Motion Directions) in one session. Observers were instructed to take a 5-minute break halfway through the session. The experimental session took about 35 minutes to complete.

3.2.2. Results

Following the data description in Moore and Enns (2004), the average proportion of trials that the observers reported alignment of the target disc and the flash was plotted for each stream condition (Figure 3.3; only data of the two baseline flash conditions and the flash condition where the target disc and the flash were physically aligned are shown).

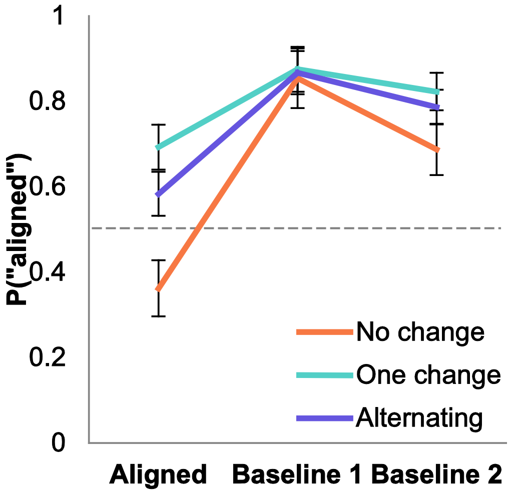

Figure 3.3. The average proportion of trials the observers reported alignment of the target disc and the flash stimuli in Experiment 3; error bars represent the standard error of the mean. Reproduced from Au and Watanabe (2013a) with permission.

An omnibus 4 × 3 × 3 (Travel Distance × Flash × Stream) repeated measures ANOVA was performed on the data. The main effect of Flash condition [$F(2,22) = 25.777$, $p < 0.001$], the main effect of Stream condition [$F(2,22) = 33.997$, $p < 0.001$], and the Flash × Stream interaction [$F(4,44) = 9.685$, $p < 0.001$] were all statistically significant; the main effect of Travel Distance was not [$F(3,33) = 2.630$, $p = 0.066$]. Specific comparisons revealed that, in the condition where the target disc and the flash were physically aligned (i.e., Aligned in Figure 3.3), the No Change condition showed a significantly lower proportion of "aligned" responses [i.e., *P*("aligned")] compared to the Alternating condition, while the One Change condition showed a significantly higher proportion of "aligned" responses compared to the Alternating condition (both at $p < 0.01$, adjusted for multiple comparisons). For the Baseline 1 condition, no significant difference was found between the three stream conditions; for the Baseline 2 condition, a significant difference in the proportion of "aligned" responses was found between the No Change vs. One Change, and between the No Change vs. Alternating conditions (both at $p < 0.01$).

The results showed that the Alternating stream exhibited some degree of FLE, which was neither as strong as in the No Change stream, nor reduced to the extent of the One Change stream. This suggests that, the regular and predictable change in surface feature of the moving disc in the Alternating stream partially impaired the motion continuity of the stream, and correspondingly reduced the magnitude of the FLE partially.

3.3. Experiment 4: random color change and the magnitude of flash-lag effect

In Experiment 4, observers' performance across the No Change, One Change, and Random streams was compared. In the Random stream, the disc color changed randomly between red and green as the disc moved along its trajectory. Because the color change in each frame was random and unpredictable, it placed some uncertainty to the observer about how the disc would behave as it moved.

3.3.1. Material and methods

3.3.1.1. Observers

Twelve new, naive observers participated in the experiment. All had given informed consent prior to the experimental session.

3.3.1.2. Stimuli and procedure

The setup was basically identical to that of Experiment 3, except that a random stream of stimulus was presented instead of the Alternating stream. In the Random stream, the color of the target changed randomly (either red or green) in each frame of its motion. Similar to Experiment 3, observers reported whether the target disc was aligned with the flash when the flash occurred. Each observer performed a total of 480 trials, with a 5-minute break in the middle of the session.

3.3.2. Results

As in Experiment 3, the average proportion of trials in which the observers reported alignment of the target disc and the flash in each stream condition was plotted (Figure 3.4).

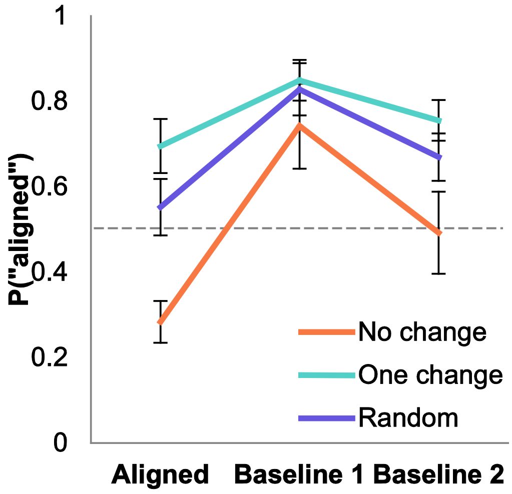

Figure 3.4. The average proportion of trials the observers reported alignment of the target disc and the flash stimuli in Experiment 4; error bars represent the standard error of the mean. Reproduced from Au and Watanabe (2013a) with permission.

An omnibus repeated measures ANOVA was performed and the results were similar to that obtained in Experiment 3. The main effect of Flash condition [*F*(2,22) = 11.581, *p* < 0.001], the main effect of Stream condition [*F*(2,22) = 14.137, *p* < 0.001], and the Flash × Stream interaction [*F*(4,44) = 6.795, *p* < 0.001] all reached significance. The main effect of Travel Distance was marginally significant [*F*(3,33) = 2.927, *p* = 0.048], and pairwise comparisons showed that the four Travel Distance conditions did not differ significantly from each other. Specific comparisons showed that when the target disc and the flash were physically aligned, the proportion of “aligned” responses was significantly lower in the No Change condition compared to the Random condition, whereas there was a significantly higher proportion of “aligned” responses in the One Change condition compared to the Random condition. Similar to Experiment 3, no significant difference was found among the three stream conditions in the Baseline 1 condition. There was a significant difference between the No Change vs. One Change, and between the No Change vs. Random stream conditions (both at $p < 0.01$) in the Baseline 2 condition.

To summarize, Experiments 3 and 4 replicated the finding that inserting a single change in the object's feature during motion (i.e., the One Change stream) eliminated (or greatly attenuated) the FLE compared to the No Change stream. Furthermore, the two experiments demonstrated that a motion stream where the object alternates colors regularly or changes color randomly elicits some degree of FLE. These results may imply that: (a) the weakened FLE in the Alternating and Random streams may be due to impaired perceptual smoothness of motion compared to the No Change stream, and (b) the elimination of FLE in the One Change stream may be due to the high salience of the target disc during the second last frame of the motion. In the Alternating and Random streams, the disc may be no longer salient at the moment of flash presentation (compared with the One Change stream) because the surface feature is continuously changing throughout the disc's motion, leading to the survival of FLE under these conditions.

3.4. Experiment 5: unexpected color change within a regular sequence of color alternations

Experiments 3 and 4 showed that salience of the color change was reduced because it was embedded in a sequence of color changes, either regularly alternating or randomly changing. If the salience of the color change was enhanced, the elimination of the FLE might become strong as the One Change stream. To test this idea, in this experiment, a new stream condition was included in which the moving disc alternated between red and green, and changed to a new color (blue) at the second last frame of the motion. It was predicted that the unexpected color change during the second last frame might restore the salience of the moving disc at the time of flash presentation, thus the FLE would be eliminated under such a condition.

3.4.1. Material and methods

3.4.1.1. Observers

Twelve new and naïve observers participated. All had normal or corrected-to-normal vision and gave informed consent prior to the experiment.

3.4.1.2. Stimuli and procedure

The stimuli and procedure were mostly identical to Experiment 3, with only one difference in the Alternating stream condition: the color of the moving disc alternated between red and green, and changed to blue for one frame during the second last frame of the motion stream (this is referred to as the "Additional Change" condition). The No Change and One Change conditions were identical to those in Experiment 3. Observers completed a total of 480 trials as in Experiment 3.

3.4.2. Results

The average proportion of trials that the observers reported alignment of the target disc and the flash were plotted for each stream condition in Figure 3.5.

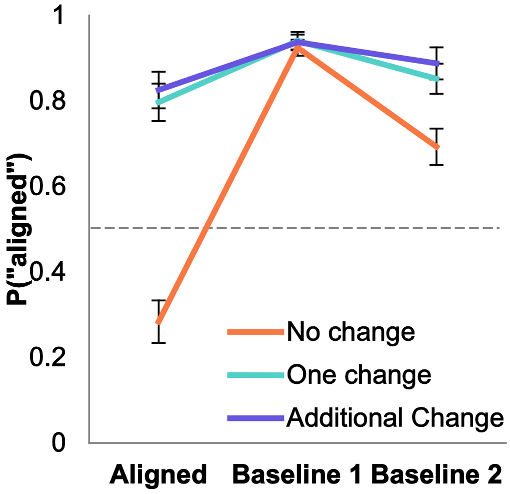

Figure 3.5. The average proportion of trials the observers reported alignment of the target disc and the flash stimuli in Experiment 5; error bars represent the standard error of the mean. Reproduced from Au and Watanabe (2013a) with permission.

Similar to the analysis in the previous experiments, an omnibus Travel Distance × Flash × Stream repeated measures ANOVA was performed on the data. The main effect of Flash condition [*F*(2,22) = 57.110, *p* < 0.001], the main effect of Stream condition [*F*(2,22) = 45.541, *p* < 0.001], and the Flash × Stream interaction [*F*(4,44) = 28.581, *p* < 0.001] were found to be significant; the main effect of Travel Distance was not [*F*(3,33) = 0.169, *p* = 0.917]. Specific comparisons showed that when the target disc and the flash were physically aligned, there was a significantly lower proportion of “aligned” responses in the No Change condition compared to the One Change and the Additional Change conditions (both at *p* < 0.01, adjusted for multiple comparisons); there was no significant difference in proportion of “aligned” responses in the One Change condition compared to the Additional Change condition. For the Baseline 1 condition, no significant difference was found among the three stream conditions. For the Baseline 2 condition, significant differences in the proportion of “aligned” responses were found between the No Change vs. One Change, and between the No Change vs. Additional Change conditions (both at $p < 0.01$).

These results showed that the introduction of an unexpected color change in the sequence of ongoing alternation of color (between red and green) eliminated the FLE, to the degree comparable to the One Change condition. Such a salient change might have strongly degraded the continuity of the object motion, enabling the visual system to correctly register the position of the moving object (which appeared in a new color), and spared the observer from perceiving the FLE.

3.5. Control experiments

As mentioned in section 3.3.2, there could be two hypotheses to explain the present pattern of results: (a) the weakened FLE in the Alternating and Random streams might be due to impaired perceptual smoothness of motion, and (b) the elimination of FLE in the One Change and Additional Change streams might be due to the high salience of the target disc during the flash presentation at the second last frame of the motion. To verify these two hypotheses, several short control experiments were conducted with five additional observers by requesting them to judge the smoothness of the motion stream or the salience of the target disc during the second last frame of the motion. In each trial of the sessions where smoothness of motion was evaluated, a No Change stream and an Alternating stream (or a Random stream in a separate session) were presented one after the other in random order, and observers indicated which of the two exhibited greater smoothness in motion. Twenty four trials were presented in each session. In most of the trials, the observers reported that the No Change stream had greater smoothness than both the Alternating and Random streams (average percentage of trials in which the No

Change stream was judged to be more smooth in comparison to the Alternating stream = 84.2%, in comparison to the Random stream = 84.2%).

The sessions for target disc salience during the second last frame of the motion were conducted in a similar way, but the One Change stream was presented instead of the No Change stream. Observers were asked to judge which of the two presented streams showed a more salient target disc at the second-last frame of the motion. They were explained explicitly that "salience" in this context refers to how strongly the moving disc stands out relative to the other moments in the motion stream. The observers judged the target disc in the One Change stream to be more salient compared to the Alternating (85.8%) and Random (90.8%) streams. The control experiments therefore suggest that both hypotheses (a) and (b) contribute to explain the reduced, but not eliminated, FLE observed in the Alternating and Random conditions.

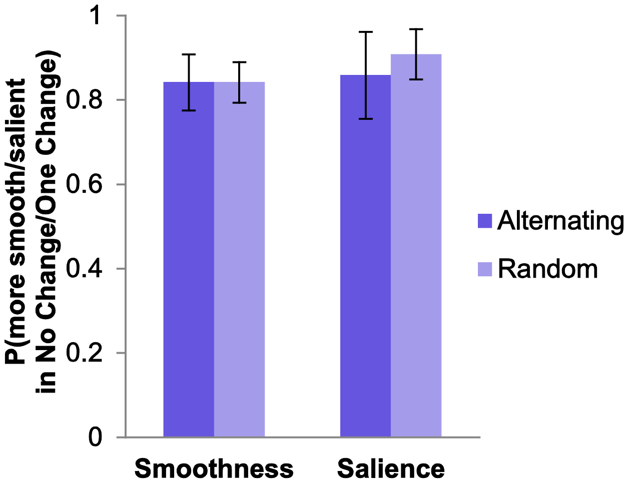

Figure 3.6. The average proportion of trials the observers reported that the motion of the No Change stream appeared to be more smooth than the Alternating/Random stream, and that the target disc looked more salient at the second last frame in the One Change stream than the Alternating/Random stream; error bars represent the standard error of the mean. Reproduced from Au and Watanabe (2013a) with permission.

3.6. Discussion

The results of the present study showed that under the conditions where the target object kept changing its color while moving in a uniform trajectory (i.e., Alternating and Random streams), a significant FLE was observed, although the magnitude was somewhat attenuated in comparison to the No Change condition. Furthermore, the results of the control experiments suggested that (a) the attenuation of the effect under those conditions might be due to the lower perceived motion smoothness compared to the No Change stream, and (b) the strong salience of the target disc at the moment when the flash occurred might be responsible for the elimination of the effect in the One Change stream. In accordance with (b), Experiment 5 showed that a sudden change to an unexpected color at the time of flash presentation in the Alternating stream (i.e., Additional Change stream) could revive the salience of the moving disc, leading to an elimination of FLE comparable to that in the One Change condition. These results suggest that smooth motion defined by unchanged surface feature is not a necessary condition for the FLE to occur. As long as the visual system identifies a single entity throughout the motion, without a salient transient change (i.e., in the cases of the Alternating and Random streams), FLE can still be observed. A highly salient change which occurs unexpectedly (i.e., in the One Change and Additional Change streams) is required to break the continuity and cause the visual system to identify the existence of multiple objects in the stream.

In the context of the FLE, the present results support the notion that spatiotemporal continuity dominates surface feature in processing object persistence (Mitroff & Alvarez, 2007). Although under some conditions, surface features can guide the mapping and updating of individual objects (Moore et al., 2010), spatiotemporal information is weighted more strongly in the computation of object persistence when both types of information are available (Tas, Dodd, & Hollingworth, 2012). A brain imaging study by Yi et al. (2008) also provided strong evidence showing that discontinued spatiotemporal trajectories can cause visually identical faces to be represented as different individual objects, in which the brain area involved was the most staunchly “featural” area of the ventral visual cortex. The determination of object persistence during object motion involves identifying the correspondence between objects separated by short instances of time. This is similar to the situation concerning how the visual system computes motion correspondence in the apparent motion phenomenon, in which solutions are sometimes needed to map multiple objects at one instance to multiple objects at other locations at another instance; in such a case, spatiotemporal information plays an important role in assisting the visual system to arrive at an appropriate solution (Dawson, 1991).

From the results of the control experiments, one can infer that observer’s perceived smoothness of object motion and salience of the transient change during motion mediate the magnitude and determine the survival of the FLE. Concerning perceived motion smoothness, the results suggest that observers’ subjective perception of smoothness was related to the magnitude of the FLE. In the Alternating and the Random conditions, the observers reported weaker motion smoothness when compared to the No Change condition, while the results of Experiments 3 and 4 indicated a significantly weaker FLE in the Alternating and Random conditions than the No Change condition. This is consistent with the previous finding that perceived motion smoothness (i.e., sampling rate of the motion trajectory) and magnitude of the FLE are correlated (Khurana, Nijhawan, & Watanabe, 1998). Such a relationship between motion smoothness and magnitude of the FLE implies that the maintenance of object files that give rise to the FLE may be associated with smoothness of motion. In the context of the present study, the rapid change in physical feature in the Alternating and Random streams impaired the perceived motion smoothness, and the maintenance of object files was thus degraded, which led to a smaller FLE. Although the maintenance process for object files was partially interrupted, the visual system might still identify only one object existed in the motion stream. For salience of the transient change at the time of flash presentation, the results are consistent with the proposal of Moore and Enns (2004) that the FLE depends on such a salient and unexpected change in smooth motion, as abrupt changes in object features may disrupt object representations (Moore et al., 2007). One possibility is that the salient and unexpected change in the One Change stream captured the observer’s attention, which breaks the continuity of the object motion as a single event. At the moment of flash onset, the abrupt change in the moving object raises attention level and allows the moving object to be associated with the flash onset at its veridical position, sparing it from the FLE. In the Alternating and Random streams, since the color change was ongoing, any change would become less salient and less able to capture attention, thus the FLE could be preserved. Furthermore, the results from Experiment 5 suggested that in the Alternating stream with a change to a new color (i.e., the Additional Change stream), the additional change to an unexpected color could lead to the elimination of the FLE which is comparable to that in the One Change stream. This confirmed the prediction that a highly salient and unexpected change strongly captures the observer’s attention and spared the FLE as mentioned above. Furthermore, the experiments of Moore and Enns (2004) suggested that an unexpected color change in a smooth motion stream might lead the visual system to interpret the scene as containing two separate objects, thus leading to the perception of seeing two objects in the scene. The results of Experiment 5 in the present study is consistent with this suggestion, that inserting an unexpected color change (blue) within the not-so-smooth stream of red-green alternations eliminated the FLE. Here, the elimination of FLE in such a condition (comparable to the One Change condition) might suggest that the blue object in the alternating stream is interpreted as a separate object from “the object” that has moved in regular trajectory with constantly alternating colors.

3.7. Conclusion

In sum, the present study extended the previous findings of FLE experiments (e.g., Moore & Enns, 2004) and showed that the FLE can occur in motion streams where the physical features of the moving object keeps changing along with the motion. The magnitude and survival of the FLE seemed to be related to perceived motion smoothness and salience of the moving object at the time of the flash. The results suggest that ongoing changes in a physical feature partially degrade the maintenance of the object file, but do not eliminate the overall percept of only one object in the motion stream. At the same time, it mostly reduces the salience of the change at the moment of flash presentation. However, if the object unexpectedly changes into a new color at the time of the flash presentation, it may make the change salient again and capture the observer’s attention, sparing the observer from perceiving the FLE. Considering the FLE as a perceptual phenomenon itself, the present results do not constraint much about the current theories of the FLE, but are informative about how changes in object features, predictability and continuity play their roles in the processing of object persistence during dynamic visual events. As discussed, this study leaves open the question regarding the role of attention at the moment of flash presentation on the resulting mislocalization effect. Future studies may focus on how attention plays its role in the occurrence of the FLE.

In this chapter, we examined how a briefly-presented static stimulus can affect a target object which dynamically moves in a constant velocity in the space. In the next chapter, we focus on how objects presented dynamically for brief instants at different positions can influence other objects at other different spatiotemporal positions.

CHAPTER 4

NUMEROSITY UNDERESTIMATION WITH ITEM SIMILARITY IN DYNAMIC VISUAL DISPLAY

In this chapter, I report a new visual illusion in which the total number of elements contained in a scene is perceived to be less than it actually is, when they are presented in a fast rate over a period of time. Using static and dynamic displays with different presentation rates and placing different levels of attentional load on the observer, the experiments investigated the situations that such underestimation of numerosity would occur. The results are then discussed with how the visual system behaves when processing rapid dynamic information presented in close spatiotemporal proximity in relation to the object substitution account.

4.1. Background and objectives

When viewing a complex scene that contains a large number of objects, humans seem to have no difficulty in approximating the numerosity. The visual system is so proficient at doing the job as if it can sense number directly (Ross & Burr, 2010). We can judge numerosity in various ways, such as one-by-one serial counting of elements, rough estimation based on instantaneous impressions, or discriminating numerosity between two scenes.

It has been suggested that the visual system may possess two separate systems for processing numerosities of different ranges (small versus large number of objects) without counting. People are quite accurate at rapidly judging the exact number of objects up to about four. This rapid and accurate judgment of numerosity for a range of small numbers without counting is referred to as “subitizing” (Dehaene, 1992; Kaufman et al., 1949). The other system deals with larger numbers of items and involves approximate estimation—the rapid, coarse approximation of numerosity which is less accurate and precise than subitizing. Evidence supporting the idea that there are two separate processing systems has been accumulating (Ansari et al., 2007; Hyde & Spelke, 2011; Palomares & Egeth, 2010; Revkin et al., 2008). However, a few studies have also shown that adaption and manipulation of attentional load could affect numerosity judgments in both high (estimation) and low (subitizing) ranges of numerosities, suggesting that numerosity may be processed by a single system (Burr, Anobile, & Turi, 2011; Vetter, Butterworth, & Bahrami, 2008). Whether numerosity judgments depend on a single process or multiple processes is still under debate.

In the current literature, many studies so far have examined numerosity estimation in vision with one-shot static displays that requires the observer to give an estimate or to compare numerosities between briefly presented stimuli. However, despite the fact that numerosity judgments in the real life often occur in dynamic contexts, very few investigations of numerosity judgments with dynamic displays have been conducted. For example, by asking observers to indicate which of the two streams of dynamic dot displays appeared more numerous, Allik and Tuulmets (1993) found that perceived numerosity decreased when spatial and temporal proximity between the presented items increased. This finding demonstrates how temporal properties of dynamic visual events can interact with spatial properties on influencing the perception of overall numerosity.

To add new understanding to the currently limited literature on numerosity judgments in dynamic displays, the present study examined whether the homogeneity of visual elements in dynamic displays would affect numerosity judgments. It was hypothesized that, in addition to spatial and temporal proximity (Allik & Tuulmets, 1993), proximity in feature space (i.e., similarity) would also affect numerosity judgments. The results of the experiments supported this hypothesis and demonstrated a new phenomenon—underestimation of numerosity in dynamic displays with a large number of objects (which is far beyond the subitizing range) when a visual feature (color) is identical among items, compared to the condition where items differ in color. Further, the study demonstrated that the underestimation effect due to color homogeneity occurred only with fast presentation rate but neither with slow presentation rate (Experiment 6) nor in static displays (Experiment 7).

Previous studies have also shown that subitizing can be affected by the availability of attentional resources (Burr, Turi, & Anobile, 2010; Egeth, Leonard, & Palomares, 2008; Olivers & Watson, 2008; Railo et al., 2008). Therefore, in Experiments 8 and 9, attentional load was manipulated by requiring observers to perform another task simultaneously with the numerosity judgment task and examined whether the availability of attentional resources would affect the underestimation effect. Finally, I discuss possible explanations for the numerosity underestimation effect, and suggest that object substitution (Enns & Di Lollo, 1997) might have occurred among items that are close on both spatiotemporal and featural dimensions, which caused the apparent perception of a smaller number of objects.

4.2. Experiment 6: perceived numerosity in displays with elements of identical versus different appearance

The aim of Experiment 6 was to examine whether the homogeneity of visual elements in a dynamic display would influence numerosity judgment. Two streams of dynamic displays were presented in succession: one stream contained visual elements in a single color and the other contained visual elements in two different colors.

4.2.1. Methods

4.2.1.1. Observers

Twelve naïve participants were recruited to participate in the experiment. All observers had normal or corrected-to-normal visual acuity. All gave informed consent prior to the experiment.

4.2.1.2. Stimuli

Stimuli were programmed in MATLAB R2012b (MathWorks, USA) with the Psychophysics Toolbox extension (version 3.0.10; Brainard, 1997; Pelli, 1997) and were displayed on a CRT monitor with a refresh rate of 100 Hz and resolution of 800 × 600 pixels. The stimulus presentation and response collection were controlled by a personal computer running with the Windows 7 operating system. Observers viewed the display at a distance of 60 cm from the monitor and performed the experiment in a dimly illuminated and quiet room.

All stimuli were presented on a black background. The fixation stimulus was a cross composed of vertical and horizontal white lines (length = 0.317°, width = 0.045°), appearing at the center of the screen. Dot stimuli were either red or green (luminance = 0.47 cd/m^2) with a diameter of 0.273°, appeared at one of 72 possible positions evenly distributed on an imaginary circle (radius = 4.533°) around the fixation.

4.2.1.3. Procedure

Each trial began with a blank screen. The observer initiated each trial by pressing the space bar on the keyboard. The fixation then appeared and stayed on the screen throughout the trial until a response was given. After 500 ms had passed, two streams of dynamic visual stimulus appeared on the screen, separated by a blank screen of 500 ms

(Figure 4.1). The two streams were always presented for 960 ms each. One of the streams consisted of dots with the same color (either all red or all green; same-color stream), while the other stream consisted of dots of both colors (different-color stream).

The experiment consisted of two blocks: one with a fast presentation rate of 40 ms per frame and the other with a slow rate of 240 ms per frame. Since the duration of each stream was fixed at 960 ms, a total of 24 frames were presented in each stream in the fast-rate (40 ms/frame) block, and four frames were presented in each stream in the slow-rate (240 ms/frame) block. In the different-color stream, two dots (one red, one green) were always presented in each frame at two randomly determined positions among the 72 possible positions, with the constraint that the two randomly selected positions in each frame must be different from the two selected in the previous frame. Thus, for different-color streams, a total of 2 dots × 24 frames = 48 dots were presented in the fast-rate block, and 2 dots × 4 frames = 8 dots were presented in the slow-rate block.

In the same-color stream, the number of dots in each frame varied: one, two, or three dots were presented in each frame. Similar to the different-color streams, the positions of the dots were constrained by the condition that the randomly selected positions did not repeat any of those presented in the previous frame. In each same-color stream, the percentage of frames that displayed one or three dots was varied, across a range of nine conditions: four conditions with 25%, 50%, 75%, or 100% of the frames showing one dot; four conditions with 25%, 50%, 75%, or 100% of the frames showing three dots; and the 0% condition in which two dots were presented in all frames in the stream. For example, in the 25% one-dot condition in the fast-rate block, 25% of the frames (which equals 24 frames × 25% = 6 frames) in the stream showed one dot, and 75% of the frames (18 frames) showed two dots. The two types of frames were presented in a random sequence within the stream. Consequently, there were nine different "number of dots" conditions showing a total of 24, 30, 36, 42, 48, 54, 60, 66, or 72 same-color dots in streams in the fast-rate block, and 4, 5, 6, 7, 8, 9, 10, 11, or 12 same-color dots in streams in the slow-rate block (both are represented in percentage of numerosity relative to the standard, which are 48 and 8, respectively, in Figure 4.2 in the Results section).

Observers were instructed to indicate which of the two streams contained a greater number of dots, by pressing one of two designated keys on the keyboard, after the second stream disappeared. For each block, observers completed a total of 180 trials (2 presentation orders × 9 percentages of frames or total number of dots in the same-color stream × 2 colors of the same-color stream × 5 repetitions). They completed the two blocks in a counter-balanced order with a short break of five minutes in between. The whole experiment took about 30 minutes to complete.

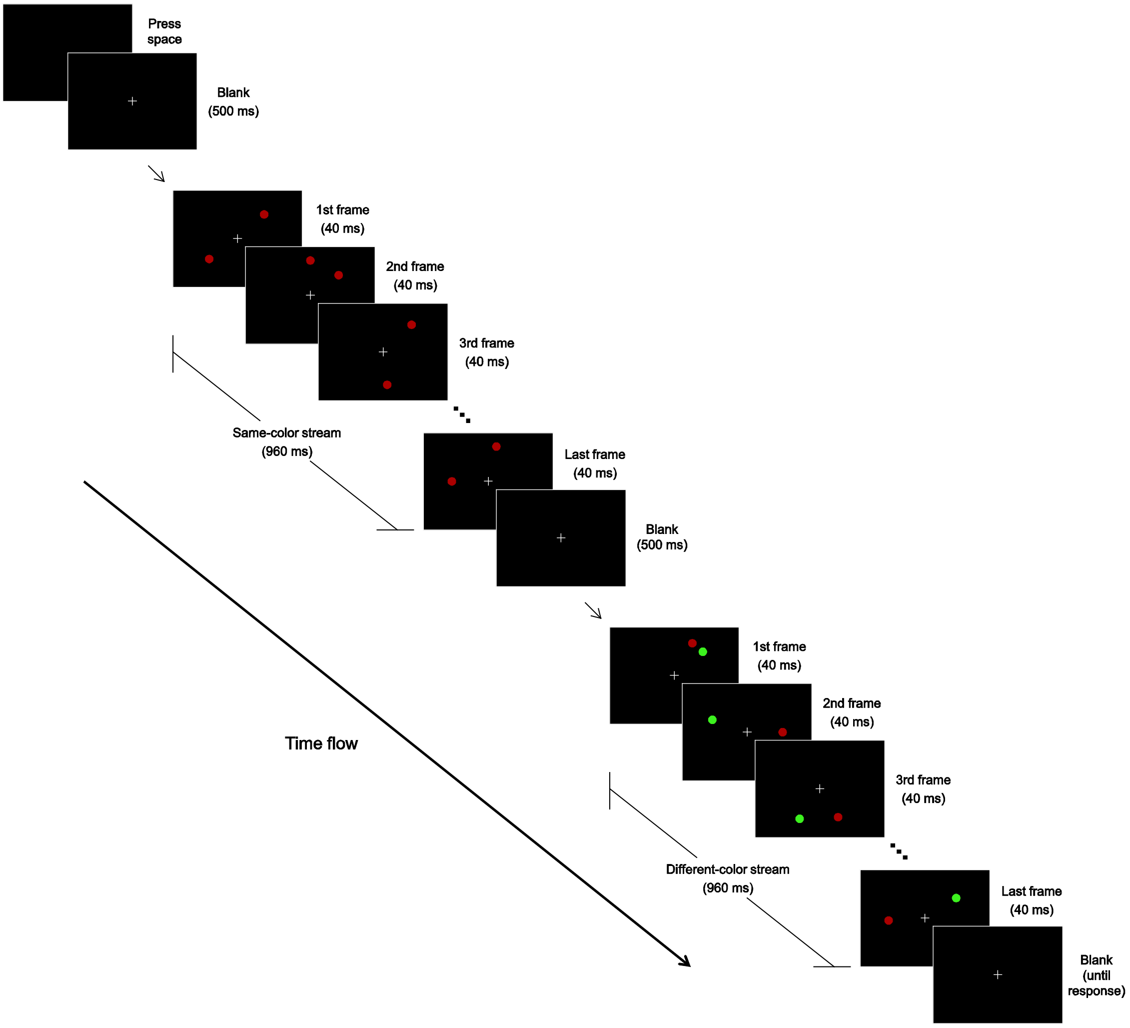

Figure 4.1. The flow of a trial in Experiment 6 at the frame duration of 40 ms. Reproduced from Au and Watanabe (2013b) with permission.

4.2.2. Results

The proportion of trials in which the observer indicated that the same-color stream contained a greater number of dots than the different-color stream was computed for each "percentage of numerosity relative to the standard" condition. The proportions averaged across the twelve observers are shown in Figure 4.2A, separately for the fast-rate (40 ms/frame) and slow-rate (240 ms/frame) blocks.

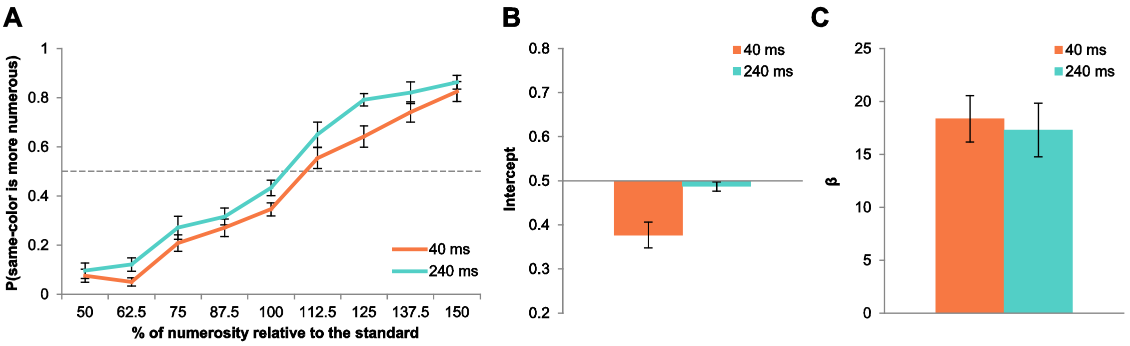

Figure 4.2. The proportion of trials the observers responded that the same-color stream contained a greater number of dots than the different-color stream in Experiment 6 (Panel A); average estimated proportion value (i.e., intercept) at $x = 100\%$ (Panel B); average β of the fitted logistic function (Panel C); error bars showing the standard error of the mean. Reproduced from Au and Watanabe (2013b) with permission.

An omnibus 2 × 2 × 9 (Order × Presentation Rate × Numerosity) ANOVA was performed to examine whether there was any difference in proportion of response across the conditions. The main effects of Presentation Rate [$F(1,11) = 13.856$, $p = 0.003$] and Numerosity [$F(8,88) = 108.293$, $p < 0.001$], the Order × Presentation Rate interaction [$F(1,11) = 10.770$, $p = 0.007$], and the Order × Numerosity interaction [$F(8,88) = 5.247$, $p < 0.001$] were found to be statistically significant; while the Order main effect [$F(1,11) = 0.442$, $p = 0.520$], the Presentation Rate × Numerosity interaction [$F(8,88) = 0.953$, $p = 0.478$], and the Order × Presentation Rate × Numerosity interaction [$F(8,88)= 0.801$, $p = 0.604$] were not. Simple main effect analyses showed that the Order main effect was only marginally significant in the slow-rate condition ($p = 0.057$) and not significant in the fast-rate condition ($p = 0.428$). Therefore, the order of presentation might be of some relevance in explaining the apparent bias in response (a proportion lower than the value of 0.5 in the 100% condition in Figure 4.2A) observed in the slow-rate condition, but the underestimation effect observed in the fast-rate condition is not likely to be explained by bias in response.

For each observer, separate logistic functions were fitted to the data in the fast-rate and slow-rate blocks respectively, using the Bootstrap Inference function provided in the Psignifit toolbox for MATLAB version 3.0 (see http://psignifit.sourceforge.net/; Fründ, Haenel, & Wichmann, 2011; Wichmann & Hill, 2001). The α (the threshold obtained after adjustment of the lower and the upper bound), β (a parameter related to the slope of the fitted function, representing the variance), γ (the miss rate, governing the lower bound of the fitted function), λ (the lapse rate, governing the upper bound of the fitted function), and the intercept (the estimated proportion at the condition of 100%) of the fitted functions were determined (Figure 4.2B, 4.2C). The parameters γ and λ were allowed to vary within the range of 0 to 0.25 in the fitting procedure. The slope of a psychometric function represents the precision in making response to a particular stimulus level: when perfect judgments are made, the slope would be infinity (β would tend to zero) in the midway of the stimulus range; when judgments become imprecise, variations in response would emerge and the function would become S shape with increasingly shallow slope at all stimulus levels (i.e., a function with increasingly large β). An intercept with a value smaller than 0.5 would represent the case that the observer chooses the same-color stimulus to be more numerous than the different-color stimulus for a less proportion of trials, and thus would indicate an underestimation of numerosity of the same-color stimulus relative to the different-color stimulus in the particular condition. One-sample *t*-tests showed that the intercept in the fast-rate condition was significantly smaller than 0.5 [$t(11) = -4.203$, $p = 0.001$], while the intercept in the slow-rate condition was not [$t(11) = -1.234$, $p = 0.243$]. A paired *t*-test showed that the intercept in the fast-rate condition was significantly lower than that in the slow-rate condition [$t(11) = -3.727$, $p = 0.003$]. Thus, the observers perceived a smaller number of dots in the same-color than in the different-color streams with the fast-rate presentation. However, this phenomenon was not evident with the slow-rate presentation. This implies that fast presentation of visual information, which challenges the spatiotemporal resolution of the visual system, is a critical factor for eliciting the underestimation effect. For the β of the fitting functions, the two conditions did not show any significant difference [$t(11) = 0.620$, $p = 0.548$].

4.2.3. Control experiments

Flickering dots can induce illusory motion in some cases. In a control experiment using the identical set up as in Experiment 6, ten observers were asked to judge which of the two presented stimulus streams showed more smooth motion. Results revealed that the observers did not perceive the same-color stream and the different-color stream differently in terms of motion smoothness, for both the fast-rate [$t(9) = 1.550$, $p = 0.156$] and the slow-rate [$t(9) = 0.563$, $p = 0.587$] conditions (testing the intercept of the curve fitted with logistic function; Figure 4.3).

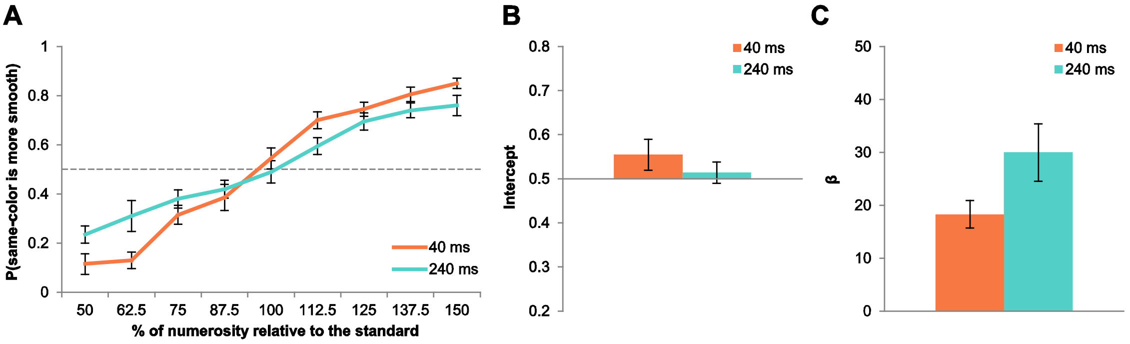

Figure 4.3. The proportion of trials the observers responded that the same-color stream showed more smooth motion than the different-color stream in the first control experiment (Panel A); average estimated proportion value (i.e., intercept) at x = 100% (Panel B); average β of the fitted logistic function (Panel C); error bars showing the standard error of the mean. Reproduced from Au and Watanabe (2013b) with permission.

In another control experiment, also using the identical setup as in Experiment 6, ten observers were asked to judge which of the stimulus streams flickered at a higher frequency (although the two movie streams were always at the same flickering rate). Observers did not perceive different flickering frequencies in the same-color and the different-color streams for both the fast-rate [$t(9) = -1.385$, $p = 0.199$] and the slow-rate [$t(9) = -1.340$, $p = 0.213$] conditions (Figure 4.4).

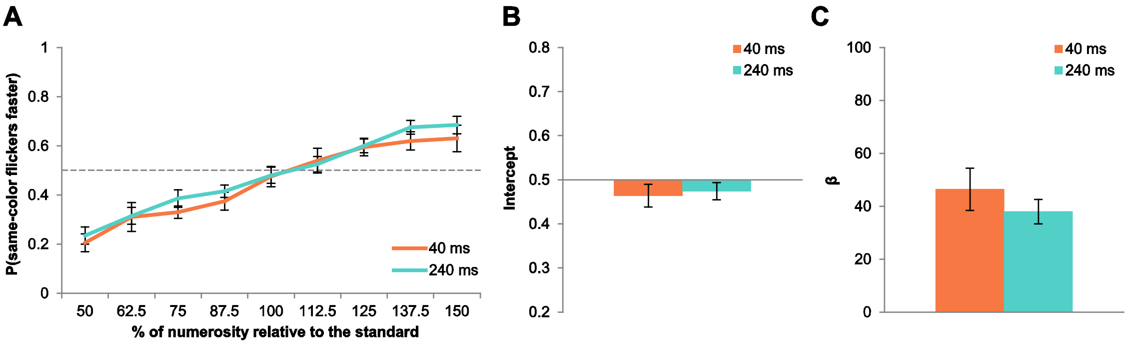

Figure 4.4. The proportion of trials the observers responded that the same-color stream flickered at a higher frequency than the different-color stream in the second control experiment (Panel A); average estimated proportion value (i.e., intercept) at x = 100% (Panel B); average β of the fitted logistic function (Panel C); error bars showing the standard error of the mean. Reproduced from Au and Watanabe (2013b) with permission.

Thus, the control experiments showed that the underestimation effect in the same-color dot streams relative to different-color dot streams was not likely to be due to differently perceived motion smoothness or flickering rate.

4.3. Experiment 7: numerosity judgments in static displays

Experiment 7 aimed to investigate whether the effect of numerosity underestimation observed in Experiment 6 occurs only in dynamic displays or whether it could be generalized to static displays as well. Observers performed a numerosity judgment task similar to Experiment 6, but the dot stimuli were presented on the screen only once for varied durations.

4.3.1. Methods

4.3.1.1. Observers

Ten new naïve observers were recruited to participate in the experiment. All had normal or corrected-to-normal vision and gave informed consent prior to the experiment.

4.3.1.2. Stimuli and procedure

The experimental flow and the stimuli were basically the same as in Experiment 6, but instead of presenting two 960-ms streams of dynamic stimuli, two static frames were presented sequentially for short (40 ms), medium (240 ms), or long (960 ms) durations, separated by a 500-ms blank interval. In the different-color stimulus, a total of 48 dots were presented in random positions selected from the 72 possible positions, with half of them in red and half of them in green. In the same-color stimulus, a total of 24, 30, 36, 42, 48, 54, 60, 66, or 72 dots were presented in random positions, which were either all in red or all in green. The experiment was conducted in three separate blocks with the three different presentation durations. Observers performed 180 trials in each of the three blocks (in counter-balanced order among observers) separated by 5-minute breaks.

4.3.2. Results

As in the dynamic display session (Experiment 6), the proportion of trials that observers responded the same-color stimulus contained more dots than the different-color stimulus was computed separately for the short duration (40 ms), medium duration (240 ms), and long duration (960 ms) blocks (Figure 4.5A).

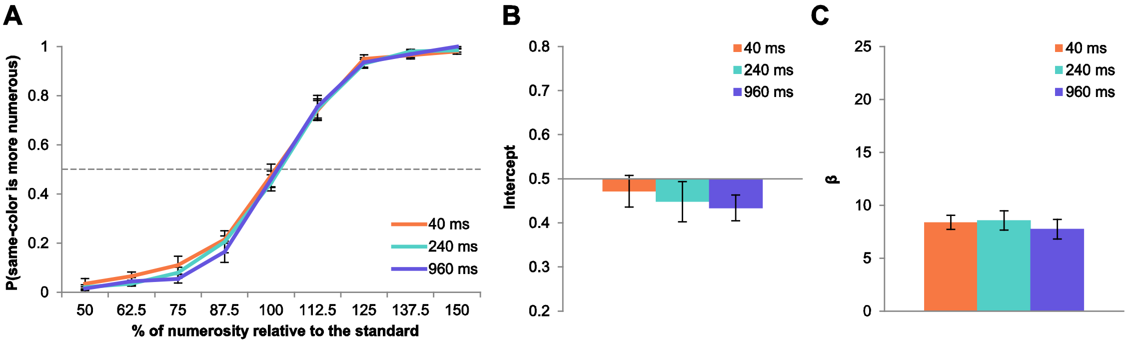

Figure 4.5. The proportion of trials the observers responded that the same-color stimulus contained a greater number of dots than the different-color stimulus in Experiment 7 (Panel A); average estimated proportion value (i.e., intercept) at x = 100% (Panel B); average β of the fitted logistic function (Panel C); error bars showing the standard error of the mean. Reproduced from Au and Watanabe (2013b) with permission.

Following the analysis in Experiment 6, an omnibus 2 × 3 × 9 (Order × Presentation Duration × Numerosity) ANOVA was performed on the data. The main effects of Order [$F(1,9) = 25.988$, $p < 0.001$] and Numerosity [$F(8,72) = 594.673$, $p < 0.001$], and the Order × Numerosity interaction [$F(8,72) = 11.149$, $p < 0.001$] were found to be significant; while the Presentation Duration main effect [$F(2,18) = 0.683$, $p = 0.518$], the Order × Presentation Duration [$F(2,18) = 0.152$, $p = 0.860$], the Presentation Duration × Numerosity [$F(16,144) = 0.390$, $p = 0.983$], and the Order × Presentation Duration × Numerosity interactions [$F(16,144) = 1.094$, $p = 0.366$] were not.

Logistic functions were fitted to each observer’s data separately for the three blocks (Figures 4.5B, 5C). One-sample *t*-tests showed that the intercepts in the 40-ms, 240-ms, and 960-ms conditions at the relative numerosity of 100% (48 dots; where the number of the dots was equal in the same-color and different-color streams) were not significantly different from 0.5 [40 ms: $t(9) = -0.791$, $p = 0.449$; 240 ms: $t(9)= -1.132$, $p = 0.287$; 960 ms: $t(9) = -2.260$, $p = 0.050$]. Furthermore, a one-way repeated measures ANOVA revealed no significant difference in estimated intercept between the three conditions [$F(2,18) = 0.252$, $p = 0.780$]. These results indicated that, in static displays, the observers did not perceive a smaller number of dots in same-color displays compared to different-color displays. For the estimated β of the fitted functions, no significant difference was found among the three conditions [$F(2,18) = 0.461$, $p = 0.638$].

Numerosity underestimation of the same-color displays occurred only with the fast-rate dynamic displays. This may be relevant to the fact that the visual system has a limit in its spatiotemporal resolution for processing incoming information. At a slow presentation speed, the visual system is still able to identify and localize each individual object with accuracy. If the incoming information enters the visual system at a speed beyond its processing capacity, some information may go into the system without being fully processed and without ultimately reaching conscious perception. This can be related to the object substitution account of visual masking proposed by Enns and Di Lollo (1997; see also Di Lollo et al., 2000), according to which objects can go unnoticed by conscious perception (in other words, objects are “masked”) if new information is rapidly fed into the processing system before the previous information has been fully processed. In such a case, the representation of the scene at the present moment overwrites and substitutes that of the previous moment. In Experiment 6, when the dots were presented at a fast rate in the dynamic condition, processing of some dots might not have been completed and were consequently masked. Spatiotemporal proximity (Allik & Tuulmets, 1993) and proximity in feature space (i.e., similarity) of same-color dots might have created the illusion of smaller numerosity due to masking among the same-color dots or registration of multiple dots as a single object by the processing system.

4.4. Experiment 8: the role of attention in numerosity judgments for dynamic display in fast-rate presentation

Studies have suggested that the processes of counting and subitizing are different, in that counting requires spatial attention while subitizing relies on a limited-capacity process that occurs before attention and after pre-attentive feature detection and grouping operations (Trick & Pylyshyn, 1993, 1994). However, recent studies that combined numerosity judgments with attentional blink or inattentional blindness tasks to manipulate the amount of available attentional resources have suggested that subitizing also requires attention. Specifically, these studies have demonstrated remarkable decreases in subitizing accuracy when observers engaged in an attentionally demanding primary task, while numerosity estimation is not affected by additional attentional demands (Burr et al., 2010; Egeth et al., 2008; Olivers & Watson, 2008; Railo et al., 2008; Vetter et al., 2008). Experiment 8 aimed to examine whether exploiting the availability of attentional resources by simultaneously engaging the observer in an additional task would influence the underestimation effect in same-color dynamic displays found in Experiment 6.

4.4.1. Methods

4.4.1.1. Observers

Ten new, naïve observers were recruited to participate in this experiment. All had normal or corrected-to-normal vision and gave informed consent before the experiment.

4.4.1.2. Stimuli and procedure

The experimental stimuli were created based on those used in Experiment 6. In the present experiment, only the fast-rate (40-ms frames) stimuli were used. Also, instead of presenting the whole range of the nine conditions as in Experiment 6, only five conditions were presented: 50% or 100% of frames showing one dot, 50% or 100% of frames showing three dots, and the 0% condition where two dots were presented in all frames.

In addition to the dot stimuli, four white digits (height = 0.68°, width = 0.54°, randomly selected from 1 to 9) were presented one after one at the center of the screen instead of the fixation cross during each stream in each trial. There were three attentional load conditions conducted in separate blocks: no-load, low-load, and high-load. In the no-load condition, observers were instructed to keep fixating at the changing digit while viewing the two dot streams and to judge which of the streams appeared to contain more dots as in Experiment 6. In the low-load condition, one of the four white digits (which were randomly selected) in each stream was shown in blue. Observers were asked to report the two blue digits (one for each stream) with the number pad first and then respond on which of the two streams contained more dots. In the high-load condition, two of the four digits were displayed in blue and observers had to report the four blue digits (two for each stream) before making the numerosity judgment. For the low-load and high-load conditions, observers were instructed to give priority to the digit task over the numerosity task and to ensure all digits were input correctly in each trial. There was a short break of five minutes between the blocks. The experiment took about 30 minutes to complete.

4.4.2. *Results*

The observers reported the blue digits with an average accuracy of 94% in the low-load condition and 88.1% in the high-load condition. The proportion of trials the observers judged the same-color stream contained a greater number of dots was computed for each of the three attentional load conditions (only included trials where all the blue digits were correctly reported). The results are shown in Figure 4.6A.

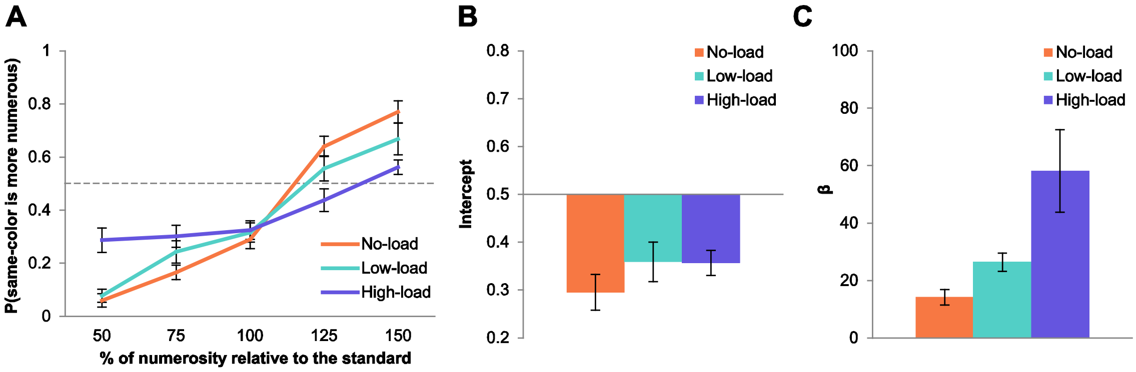

Figure 4.6. The proportion of trials the observers responded that the same-color stream contained a greater number of dots than the different-color stream in Experiment 8 (Panel A); average estimated proportion value (i.e., intercept) at x = 100% (Panel B); average β of the fitted logistic function (Panel C); error bars showing the standard error of the mean. Reproduced from Au and Watanabe (2013b) with permission.

An omnibus 2 × 3 × 5 (Order × Attentional Load × Numerosity) ANOVA was performed, and the results revealed a significant main effect of Numerosity [$F(4,36) = 103.456$, $p < 0.001$], Order × Attentional Load interaction [$F(2,18) = 7.761$, $p = 0.004$], and Attentional Load × Numerosity interaction [$F(8,72) = 8.384$, $p < 0.001$]. The main effects of Order [$F(1,9) = 3.751$, $p = 0.085$] and Attentional Load [$F(2,18) = 0.121$, $p = 0.887$], the Order × Numerosity interaction [$F(4,36) = 2.132$, $p = 0.097$], and the Order

× Attentional Load × Numerosity interaction [$F(8,72) = 0.456$, $p = 0.883$] were not significant. Simple main effect analyses revealed a significant main effect of Order in the high-load condition only ($p < 0.05$), suggesting that the presentation order might have some influence on observers' response when the task became difficult in the high-load condition. In addition, the main effect of Numerosity was significant in all the three load conditions (all $p < 0.001$).

As Experiment 6 established the underestimation effect for the same-color stream, the same logistic fitting was performed and one-tailed hypothesis tests were employed to examine whether the intercepts were smaller than 0.5 (Figure 4.6B, 4.6C). One-sample one-tailed *t*-tests revealed that the intercepts were all significantly smaller than 0.5 in the no-load, low-load, and high-load conditions [no-load: $t(9) = -5.465$, $p < 0.001$; low-load: $t(9) = -3.401$, $p = 0.004$; high-load: $t(9) = -5.521$, $p < 0.001$]. A one-way repeated measures ANOVA revealed no significant difference in intercept among the three conditions [$F(2,18) = 0.771$, $p = 0.477$]. For the β of the fitted functions, the three load conditions differed significantly [$F(2,18) = 6.840$, $p = 0.006$], with the steepest slope in the no-load condition and the shallowest slope in the high-load condition. The β of the no-load vs. high-load pair and low-load vs. high-load pair were significantly different from each other ($p < 0.05$ corrected for multiple comparisons), while the no-load vs. low-load pair was not.

The comparable intercepts in all three conditions (smaller than 0.5 and not significantly different from each other) suggest that comparable numerosity underestimation effects exhibited in all conditions. In contrast, the significant difference in slopes between the conditions suggest that as the attentional load (i.e., the task difficulty) increased, the precision of numerosity judgments decreased, with the shallowest slope indicating the lowest precision in the high-load condition.

The present results regarding the effect of attentional load can be related to several previous findings. Egeth et al. (2008) used the rapid serial visual presentation paradigm to present streams of letters, among which contained a target letter (in a specified color), to study the effect of manipulating attention on numerosity judgments. Observers had to correctly report the target letter and at the same time judge the number of dots presented in the peripheral region of the screen. It was found that when the dots were presented soon after the presentation of the target letter (i.e., the period of “attentional blink”), the performance on the numerosity task was markedly reduced, even when the number of objects was clearly within the subitizing range (e.g., two or three objects). Such results were also supported by evidence from an event-related potential (ERP) study (Xu & Liu, 2008). In the study by Olivers and Watson (2008), the lag between the letter identification task and the numerosity judgment task influenced the performance, which powerfully demonstrated that attention is involved in numerosity judgments in the subitizing range. In addition to using attentional blink tasks, dual-task experiments that control the amount of available attentional resources by engaging the observer in a spatial attention task (Burr et al., 2010, 2011) also suggested that an attention-dependent mechanism is responsible for subitizing but not estimation of larger numbers. The numerosity judgments in our task with dynamic displays was in the range of estimation (not subitizing), which has been found to be unaffected by attentional load in previous studies using static displays (Burr et al., 2010). Burr et al. (2010, 2011) suggested that an increase in attentional load decreased precision (i.e., increased variability in responses) of numerosity judgments while accuracy (which reflects the mean perceived numerosity) was not affected. This is consistent with the present results that the underestimation effect (based on mean perceived numerosity) was evident regardless of the condition of attentional load, while attentional load increased the variability in the responses (resulting in the shallower slopes of the response curves in the low-load and high-load conditions).

4.5. Experiment 9: the role of attention in numerosity judgments for static display

Experiment 9 was conducted as a control experiment and was the static display version of Experiment 8. This experiment examined whether attentional availability would interact with numerosity estimation for static displays.

4.5.1. Methods

4.5.1.1. Observers

Ten new naïve observers were recruited to participate in Experiment 9. All had normal or corrected-to-normal vision and gave informed consent before the experiment.

4.5.1.2. Stimuli and procedure

The stimuli and procedure were based on Experiment 7, but only the 960-ms stimuli were used. Furthermore, instead of presenting the whole range of number of dots conditions (the nine levels from 24 to 72 dots), only the conditions with 24, 36, 48, 60, and 72 dots were presented. The three attentional load conditions and the dual-task design followed Experiment 8. For each trial in the low-load and high-load blocks, observers first reported the digits in blue using the number pad and then judged whether the first or the second display contained a greater number of dots. The three attentional load blocks (no-load, low-load, and high-load) were conducted in counter-balanced order, with a 5-minute break between the blocks.

4.5.2. Results

The observers reported the digits with an average accuracy of 93.5% in the low-load condition and 84.4% in the high-load condition. The data of numerosity judgment are shown in Figure 4.7A.

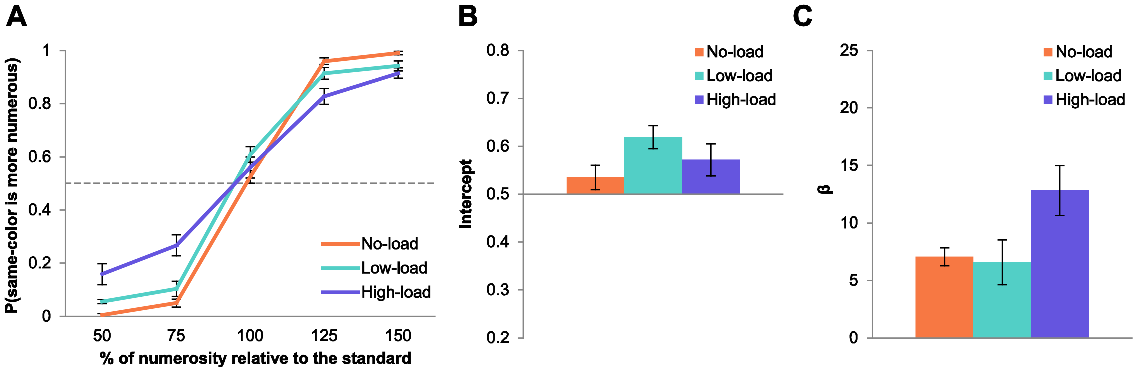

Figure 4.7. The proportion of trials the observers responded that the same-color stimulus contained a greater number of dots than the different-color stimulus in Experiment 9 (Panel A); average estimated proportion value (i.e., intercept) at x = 100% (Panel B); average β of the fitted logistic function (Panel C); error bars showing the standard error of the mean. Reproduced from Au and Watanabe (2013b) with permission.

An omnibus 2 × 3 × 5 (Order × Attentional Load × Numerosity) ANOVA showed significant main effects of Order [$F(1,9) = 20.995$, $p = 0.001$] and Numerosity [$F(4,36) = 861.909$, $p < 0.001$], Order × Numerosity interaction [$F(4,36) = 6.878$, $p < 0.001$], and Attentional Load × Numerosity interaction [$F(8,72) = 10.307$, $p < 0.001$]. The Attentional Load main effect [$F(2,18) = 3.245$, $p = 0.063$], the Order × Attentional Load Interaction [$F(2,18) = 2.024$, $p = 0.161$], and the Order × Attentional Load × Numerosity interaction [$F(8,72) = 0.412$, $p = 0.910$] were found to be not significant.

Simple main effect analyses found that the effect of Numerosity was significant across all the three load conditions (all $p < 0.001$).

In order to test whether the intercepts estimated at the numerosity value of 48 dots in each load condition were significantly smaller than 0.5, one-sample *t*-tests were conducted. The analysis showed that the intercepts for all the three conditions were not significantly smaller than 0.5 [no-load: $t(9) = 1.368$, $p = 0.898$; low-load: $t(9) = 4.997$, $p = 0.999$; high-load: $t(9) = 2.134$, $p = 0.969$]. A one-way repeated measures ANOVA found no significant difference in intercepts across the three load conditions [$F(2,18) = 2.969$, $p = 0.077$]. The data is therefore in line with those obtained from Experiment 7 and confirms that the underestimation effect does not occur in static displays. There was a marginally significant difference in β of the fitted function between the three conditions [$F(2,18) = 3.436$, $p = 0.054$], with the shallowest slope in the high-load condition. This again might reflect the effect of attentional load on the precision of numerosity judgment.

4.6. Discussion

The present study demonstrated numerosity underestimation when a large number of same-color objects are displayed in a dynamic stream, compared with different-color objects in a similar configuration. This underestimation effect was only observed in fast-rate dynamic streams, but not in slow-rate streams or static displays. This might be due to the occurrence of object substitution (Di Lollo et al., 2000; Enns & Di Lollo, 1997) among same-color objects that are spatiotemporally and featurally proximal during the high-speed presentation of stimuli.

Putting the idea of object substitution (introduced in Chapter 1) to the context of the present study, in the fast-rate same-color streams, as new dots keep appearing at different locations around the screen, new incoming information keeps congesting the input layer of the processing system. Information about a dot at a previous moment is quickly replaced by new dots at new positions without being adequately processed to enter conscious perception. If the position of a dot is close to a subsequently presented dot, the new dot is more likely to substitute the old dot without letting the old dot to reach consciousness, leading to an overall perception of smaller numerosity. This is consistent with the prediction made by Allik and Tuulmets (1993) that perceived numerosity decreases with increased spatial and temporal proximity between items in the scene. Furthermore, object substitution (or object updating) has been found to be eliminated when the salience of the critical event is high (Moore & Enns, 2004). As reviewed in Chapter 3, the FLE effect was eliminated when the tracked moving object was made salient by abruptly changing its color at the moment of the flash presentation. This might explain the difference in perceived numerosity between the different-color streams and the same-color streams. Since the different-color streams are composed of a mix of different dots and have greater color contrast than the same-color streams, the different color dots might appear more salient to the observer, which might have strengthened the signals at the pattern layer of the system. The stronger signal allows the information to remain in the pattern layer for a longer period before it completely decays. This increases the chance that the information can be fully processed by the system and reach conscious perception. As a result, object substitution in the different-color condition does not occur as strongly as in the same-color condition.

Other models concerning perceived numerosity, such as the occupancy model proposed by Allik and Tuulmets (1991), may also offer relevant explanations to the effects observed here. The occupancy model describes that, in a field containing a large number of dots, each dot occupies a circular territory around itself, and the visual system judges the numerosity based on the total size of the area occupied by all the dots. In the case when the territories of two dots overlap, the total area occupied by all the dots will be smaller. Under such a case, the observer would tend to perceive a smaller numerosity compared to the case without overlapping of dot territories (i.e., dots that are distributed more evenly and sparsely) with the same number of dots. In other words, two closer dots have a smaller total impact on perceived numerosity than two dots that are further away from each other: a dot is masked by another nearby dot (Allik & Tuulmets, 1991). In the dynamic display condition in Experiment 6, the different-color stream might have greater perceived contrast than the same-color stream and therefore may appear more salient to the observer. Such increased salience might minimize the effect of masking among dots at neighboring positions, leading the visual system to clearly regard them as two distinct identities. Given this advantage, the dot pairs of different colors can survive even when their territories overlap. This might represent an interaction between distinctive object features of the items (physical) and the occupancy of territories of each item (psychological).

He et al. (2009) showed that numerosity judgment in a cloud of dots depends on connectedness among the elements. They claimed that their observation of underestimation of numerosity in displays with connected dots reflects the processing of perceptual organization from which representations of distinct objects are formed. Specifically, the configuration of two dots joined by a line is represented as a single object as a whole, whereas the configuration with the same two dots without being connected by a line would be represented as two distinct objects. As a result, a display with lines connecting the dots would lead to the illusion of a smaller numerosity than one without connecting lines. Such perceptual organization of “single-objectness” might also be relevant in explaining the effect of numerosity underestimation in dynamic displays with elements of the same color in the present study. In the same-color dynamic displays in Experiment 6, there may exist some perceptual interactions among dots of the same color. Although the positions between two dots might be far, such interaction might lead the visual system to organize the pair of dots into a single object, given the very brief presentation duration (40 ms) in the fast-rate condition for registration of perceptual information. On the other hand, in displays showing dots of different colors, the strength of such kind of organization is much weaker as the different colors resist the tendency of the processing system in pairing up the dots. Consequently, the display with dots of the same color would be perceived as less numerous than the one with dots of different colors in the fast-rate dynamic presentation condition. This may offer an alternative explanation of the results in addition to the object substitution approach.

Experiment 8 demonstrated that the numerosity underestimation effect in dynamic displays is not affected by attentional manipulations. Although observers' precision in performing the numerosity judgment task remarkably dropped as attentional load increased (indicated by the shallower slope under the higher load conditions), the underestimation effect remained evident in all the no-load, low-load, and high-load conditions. This is consistent with previous findings showing that while the process of subitizing requires attention (Egeth et al., 2008; Olivers & Watson, 2008; Railo et al., 2008; Xu & Liu, 2008), the process of estimating large numerosities does not (Burr et al., 2010; Vetter et al., 2008). Future studies may explore the involvement of attention in numerosity judgments for dynamic stimuli across time in the subitizing range.

4.7. Conclusion

The present study demonstrated that numerosity underestimation occurs for same-color objects, compared to different-color objects, in dynamic visual displays with a fast presentation rate. This effect was not evident when a slow presentation rate or static displays were used. Using a dual-task paradigm which manipulated the attentional resources available for numerosity judgment, a high attentional load was found to produce a reduction in numerosity judgment precision; however, the underestimation effect in the same-color stream still survived. These findings are consistent with the notion of object substitution, where objects that are identical in appearance and spatiotemporally proximal may mask and substitute each other, leading to an overall perception of being less numerous in dynamic streams. Future studies may examine the effect in more natural aspects such as similarity in natural features of faces or objects in everyday life.

This chapter investigated how objects presented dynamically can mask other objects presented at different spatiotemporal positions, leading to an overall perception of a smaller numerosity. In the next chapter, I summarize the findings from the three studies described in the present thesis and conclude with their implications regarding the processing of dynamic information by the human visual system.

CHAPTER 5

CONCLUSION

5.1. Summary and theoretical implications from the present studies

This section summarizes the main and important implications that can be drawn from the present series of studies regarding perceptual processing of dynamic visual information. At a descriptive level, the present studies shed light on visual processing in different scenarios of dynamic situations.

In Chapter 2, I examined visual processing in the case of a relatively expected sequence of visual events. The experiments on the attentional repulsion and attraction effects produced results that are consistent with the previously proposed hypothesis that visual processing is postdictive and takes into account input of information within a brief time window from the target event (Eagleman & Sejnowski, 2000, 2003; as described in Chapter 1, this hypothesis is also relevant in explaining the occurrence of the FLE). The data showed that cues presented 200 ms before or after the target event could both influence the perceived position of the target, leading to mislocalization toward opposite directions. Furthermore, the condition in which the cue was invisible produced a mislocalization of the target in the same directions as in the visible condition. This is a particularly important result, because it shows that attention, as one of the core components of the human cognitive and perceptual systems, is processed below the conscious level and does not require visual awareness in order to exert influences on conscious perception. In the context of perceptual illusions, the present results represent an interesting demonstration showing that something we are not aware of can exert influence on what we are aware of.

The study on the FLE in Chapter 3 focused on visual processing in the case of an unexpected event (flash) occurring during a smooth ongoing visual event (an object moving at a constant speed). The examination of the types of visual changes that determine the magnitude of the FLE brought new understanding about the factors that are involved in the persistence and maintenance of object identity information during dynamic visual events. The experiments showed that regular alternations of object color during motion could impair the maintenance of object identity information (shown as an attenuation of FLE magnitude), but that does not eliminate or change the representation that only one object existed in the motion stream. The insertion of an unexpected color change within regular alternations between two colors, however, leads the visual system to identify multiple objects. The data also showed that the processing of visual persistence is related to and determined by the perceived smoothness of the visual event, as well as the salience of the visual object at the moment when it is affected by another unexpected event (i.e., the flash). These results highlight the key role of object features on both physical and perceptual dimensions in the visual processing of objects.

Chapter 4 examined visual processing in the case of rapid and continually-changing visual events. The dynamic presentation of visual stimuli in Experiment 6 involved a novel technique of presenting a varied number of objects in varied proportions of frames within a movie stream. This allowed the investigation of processing involved in the integration of information over a period of time by the visual system. Perceived numerosity was decreased when objects with identical visual features were presented at a fast rate, but not when presented at a slow rate or when presented in one shot. This supported the notion that the effect of substitution masking is strong when objects appear in close spatiotemporal positions (i.e., appear at close positions in space and time). In addition, engaging the observer with a secondary task to deprive his/her attention from concentrating on the numerosity judgment task affected the precision of performing the task, but did not affect the magnitude of the underestimation effect. This shows the robustness of the underestimation effect across different attentional load conditions. The results point to the importance of physical qualities (e.g., proximity in space and time, similarity in physical features) on psychological perception.

One neural model that is central to the series of studies covered in this thesis is the model of reentrant processing for visual awareness (Lamme, 2003) described in detail in Chapter 2. This model describes that as visual processing progresses from the lower primary areas to the higher areas of the brain, recurrent connections start to form between the activated neurons; visual awareness emerges when the recurrent signal can propagate through the connections up to the higher frontal areas. Attention, for example, can produce a “biased” ready state on neurons in certain brain regions so that recurrent connections can be formed more efficiently. These neurons override those in the other regions and prevent them from forming an extensive network of recurrent connections. Consequently, the information processed by the overridden neurons can hardly reach the higher areas and be consciously perceived.

This framework of recurrent processing can serve as a unified explanation applicable to the three main phenomena of visual distortion considered in the present thesis. For the attentional repulsion and attraction effects, as discussed in Chapter 2, the cue was made invisible (i.e., failed to reach visual awareness) by fast presentation of the mask in the same position of the cue upon its offset. Here, the information about the cue was quickly replaced by the mask, which covered the same physical location, without being able to form extended recurrent connections in order to be consciously perceived. In Chapters 3 and 4, objects were substituted by other objects when the processing of an earlier object is interrupted and replaced by a later object. In this situation, recurrent connections for the earlier object cannot be formed and extended to the higher cortex; thus, the representation of the object quickly fades and is erased from the observer's conscious awareness. The model offers clear distinctions that describe the fate of attended/unattended and masked/unmasked stimuli (Lamme, 2010). With fast-growing accumulation of experimental evidence supporting it (e.g. Boehler et al., 2008; Camprodon et al., 2010; Lamme et al., 2002; Pascual-Leone & Walsh, 2001; Silvanto et al., 2005; Supèr et al., 2001), this model will continue to be a good reference point for researchers studying masking and visual awareness to devise new experiments to further explore consciousness and various visual phenomena.

5.2. Implications on everyday situations and possible practical applications

Although the experiments described in the present studies examined questions that are theoretical in nature, the studies also have implications for everyday situations and applications in a practical sense. For instance, the FLE considered in Chapter 3 has been suggested to be related to offside misjudgments given by referees in football matches (Baldo, Ranvand, & Morya, 2002; Helsen, Gilis, & Weston, 2006); specifically, the kick of the attacking team is thought to initiate a sudden event (analogous to a flash) which leads the referee to perceive another player of the attacking team to be at an offside position. In a wider context, the FLE may also carry implications suggesting that a sudden unexpected change in road conditions (e.g., pedestrians crossing the road without following traffic lights) can cause misjudgments of distance between cars and pedestrians in car drivers. The numerosity underestimation effect discussed in Chapter 4 may imply that a group of people wearing the same uniform in a school, for instance, may lead to an illusion of seeing less people than there actually are.

A range of examples can be considered for practical applications too. First, understanding visual distortions can help referees in sports such as football and tennis to give fair judgments about positions in the game. During a flight, the process of landing is particularly important and accidents are frequently associated with landing. The present studies may have implications concerning this dilemma: during actual situations of landing, flashing signals around the landing path which guide the pilot may lead the pilot to misjudge positions. Pilots should therefore pay extra attention with this in mind. Furthermore, the present studies also suggest that different visual distortions of positions may be in effect according to the presentation method of landing signals: the sudden onset of an alarm when the airplane slips outside the expected landing path may cause misjudgments of position due to the FLE and the attentional repulsion and attraction effects. Although the present studies are qualitative in nature, in order to relate more tightly to real life applications, future studies can examine, quantitatively, the magnitude of visual distortion that will be in effect given certain parameters. For example, studies can be designed to measure the magnitude of the judgment error due to the FLE, with a flashing signal at a given distance. With such data, it might be possible to predict the magnitude of the distance judgment error in car drivers, when a flashing target is at a distance of ten meters, for example. With better understanding of situations that are likely to lead to distorted perceptions, we can also become more cautious about avoiding them in everyday life.

Illusions produced in laboratory settings are interesting in nature. They allow us to personally experience our own limits as humans. The human mind is a difficult subject to study, since the mechanisms of the mind are implicit. Through systematically studying illusions, we can gain knowledge about how our mind works without using invasive methods. The above is the meaning of the present thesis for the scientific community. The understanding of visual distortion and illusions also has significant meaning to normal people and to the society. This could be related to situations where judgment errors cannot be allowed. For example, in medical operations, a judgment error of even a millimeter in a surgeon could be fatal for patients. For astronauts working outside the space shuttle, any misjudgment of distance could prevent them from returning to the shuttle or to the designated orbit easily. In lawsuit cases, eyewitnesses report on what they saw about the crime which usually occurs in just a fraction of a second; as instant judgments are prone to visual distortions, professionals working in the legal field should not overlook the possibility that visual distortions are in effect in these cases. All these examples illustrate how the topic of this thesis could be related to the society. In psychophysical experiments, effects of visual distortion usually diminish as the observer repeatedly performs the same experiment. In conjunction with advanced virtual reality technology, such “practice effect” may be employed to train pilots and drivers, and any other people who cannot afford judgment errors, to adapt to, and become less prone to the influence of visual illusions.

5.3. Concluding remarks

Using psychophysical procedures, the present series of studies provided new behavioral data that revealed a number of important properties of visual information processing by the human brain. In cognitive neuroscience, direct measurements of neural activities, such as functional neuroimaging (e.g., fMRI, MEG) and electrophysiology (e.g., ERP, single-cell recording), certainly provide important data to visualize how the brain works. However, well-executed psychophysical studies can also provide indications about the mechanism underlying a perceptual phenomenon and yield interesting hypotheses and directions for further examination and validation by various neuroscientific methods. To this point, I believe the present thesis has advanced our knowledge regarding how the human visual system integrates and processes information in the dynamic visual world.

REFERENCES

Alais, D., & Burr, D. C. (2003). The “flash-lag” effect occurs in audition and cross-modally. *Current Biology, 13,* 59–63.

Allik, J., & Tuulmets, T. (1991). Occupancy model of perceived numerosity. *Perception and Psychophysics, 49*, 303–314.

Allik, J., & Tuulmets, T. (1993). Perceived numerosity of spatiotemporal events. *Perception and Psychophysics, 53,* 450–459.

Alpern, M. (1953). Metacontrast. *Journal of the Optical Society of America, 43,* 648–657.

Ansari, D., Lyons, I. M., van Eimeren, L., & Xu, F. (2007). Linking visual attention and number processing in the brain: The role of the temporoparietal junction in small and large symbolic and nonsymbolic number comparison. *Journal of Cognitive Neuroscience, 19*, 1845–1853.

Astle, D. E., Nobre, A. C., & Scerif, G. (2010). Subliminally presented and stored objects capture spatial attention. *Journal of Neuroscience, 30*, 3567–3571.

Au, R. K. C., Ono, F., & Watanabe, K. (2013a). Spatial distortion induced by imperceptible visual stimuli. *Consciousness and Cognition, 22,* 99–110.

Au, R. K. C., Ono, F., & Watanabe, K. (2013b). Retrospective perceptual distortion of position representation does not lead to delayed localization. *Advances in Cognitive Psychology, 9,* 20–31.

Au, R. K. C., & Watanabe, K. (2013a). Object motion continuity and the flash-lag effect. *Vision Research*, *92,* 19–25.

Au, R. K. C., & Watanabe, K. (2013b). Numerosity underestimation with item similarity in dynamic visual display. J*ournal of Vision, 13*(8):5. 1–15, http://www.journalofvision.org/content/13/8/5, doi:10.1167/13.8.5.

Bachmann, T., & Allik, J. (1976). Integration and interruption in the masking of form by form. *Perception, 5*, 79–97.

Baldo, M. V. C., & Klein, S. A. (1995). Extrapolation or attention shift? *Nature, 378*, 565–566.

Baldo, M. V. C., Ranvaud, R. D., & Morya, E. (2002). Flag errors in soccer games: The flash-lag effect brought to real life. *Perception, 31*, 1205–1210.

Blanco, M. J., & Soto, D. (2009). Unconscious perception of a flash can trigger line motion illusion. *Experimental Brain Research, 192*, 605–613.

Block, N. (2005). Two neural correlates of consciousness. *Trends in Cognitive Sciences, 9*, 46–52.

Boehler, C. N., Schoenfeld, M. A., Heinze, H. J., & Hopf, J. M. (2008). Rapid recurrent processing gates awareness in primary visual cortex. *Proceedings of the National Academy of Sciences of the United States of America, 105*, 8742–8747.

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

Breitmeyer, B. G. (2007). Visual masking: Past accomplishments, present status, future developments. *Advances in Cognitive Psychology, 3*, 9–20.

Breitmeyer, B. G., Kafaligönül, H., Öğmen, H., Mardon, L., Todd, S., & Ziegler, R. (2006). Para- and metacontrast masking reveal different effects on brightness and contour visibility. *Vision Research, 46*, 2645–2658.

Bridgeman, B. (1971). Metacontrast and lateral inhibition. *Psychological Review, 78*, 528–539.

Broadbent, D. E. (1958). *Perception and communication.* London: Pergamon Press.

Bullier, J., McCourt, M. E., & Henry, G. H. (1988). Physiological studies on the feedback connection to the striate cortex from cortical areas 18 and 19 of the cat. *Experimental Brain Research, 70*, 90–98.

Burr, D. C., Anobile, G., & Turi, M. (2011). Adaptation affects both high and low (subitized) numbers under conditions of high attentional load. *Seeing and Perceiving, 24*, 141–150.

Burr, D. C., Turi, M., & Anobile, G. (2010). Subitizing but not estimation of numerosity requires attentional resources. *Journal of Vision, 10*(6):20, 1–10, http://www.journalofvision.org/content/10/6/20, doi:10.1167/10.6.20.

Camprodon, J. A., Zohary, E., Brodbeck, V., & Pascual-Leone, A. (2010). Two phases of V1 activity for visual recognition of natural images. *Journal of Cognitive Neuroscience. 22*, 1262–1269.

Carrasco, M., Ling, S., & Read, S. (2004). Attention alters appearance. *Nature Neuroscience, 7,* 308–313.

Chien, S.-E., Ono, F., & Watanabe, K. (2011). Mislocalization of visual stimuli: Independent effects of static and dynamic attention. PLoS *ONE, 6*(12): e28371, 1–8, doi:10.1371/journal.pone.0028371.

Chou, W.-L., & Yeh, S.-L. (2011). Subliminal spatial cues capture attention and strengthen between-object link. *Consciousness and Cognition, 20*, 1265–1271.

Corbetta, M., & Shulman, G. L. (2002). Control of goal directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience, 3,* 201–215.

Dawson, M. R. W. (1991). The how and why of what went where in apparent motion: Modeling solutions to the motion correspondence problem. *Psychological Review, 98*, 569–603.

Dehaene, S. (1992). Varieties of numerical abilities. *Cognition, 44*, 1–42.

Deutsch, D. (2009). Auditory illusions. In E. B. Goldstein (Ed.), *Encyclopedia of perception, volume 1* (pp. 160–164). New York: Sage Publications.

Deutsch, J. A., & Deutsch, D. (1963). Attention: Some theoretical considerations. *Psychological Review, 93*, 283–321.

Di Lollo, V., Enns, J. T., & Rensink, R. A. (2000). Competition for consciousness among visual events: The psychophysics of reentrant visual processes. *Journal of Experimental Psychology: General, 129*, 481–507.

Dux, P. E., Visser, T. A. W., Goodhew, S. C., & Lipp, O. V. (2010). Delayed reentrant processing impairs visual awareness: An object substitution masking study. *Psychological Science, 21*, 1242–1247.

Eagleman, D. M. (2008). Human time perception and its illusions. *Current Opinion in Neurobiology, 18*, 131–136.

Eagleman, D. M., & Sejnowski, T. J. (2000). Motion integration and postdiction in visual awareness. *Science, 287*, 2036–2038.

Eagleman, D. M., & Sejnowski, T. J. (2003). The line-motion illusion can be reversed by motion signals after the line disappears. *Perception, 32*, 963–968.

Eagleman, D. M., & Sejnowski, T. J. (2007). Motion signals bias localization judgments: A unified explanation for the flash-lag, flash-drag, flash-jump, and Frohlich illusions. *Journal of Vision, 7*(4):3, 1–12, http://journalofvision.org/7/4/3/, doi:10.1167/7.4.3.

Egeth, H. E., Leonard, C. J., & Palomares, M. (2008). The role of attention in subitizing: Is the magical number 1? *Visual Cognition, 16*, 463–473.

Enns, J. T., & Di Lollo, V. (1997). Object substitution: A new form of masking in unattended visual locations. *Psychological Science, 8*, 135–139.

Enns, J. T., & Di Lollo, V. (2000). What's new in visual masking? *Trends in Cognitive Sciences, 4,* 345–352.

Felleman, D. J., & Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex, 1,* 1–47.

Fox, M. D., Corbetta, M., Snyder, A. Z., Vincent, J. L., & Raichle, M. E. (2006). Spontaneous neuronal activity distinguishes human dorsal and ventral attention systems. *Proceedings of the National Academy of Sciences of the United States of America, 103,* 10046–10051.

Fründ, I., Haenel, N. V., & Wichmann, F. A. (2011). Inference for psychometric functions in the presence of nonstationary behavior. *Journal of Vision, 11*(6):16, 1–19, http://www.journalofvision.org/content/11/6/16, doi:10.1167/11.6.16.

Fuller, S., & Carrasco, M. (2006). Exogenous attention and color perception: Performance and appearance of saturation and hue. *Vision Research, 46,* 4032–4047.

Gobell, J., & Carrasco, M. (2005). Attention alters the appearance of spatial frequency and gap size. *Psychological Science, 16,* 644–651.

Gordon, R. D., & Irwin, D. E. (1996). What's in an object file? Evidence from priming studies. *Perception and Psychophysics, 58,* 1260–1277.

Gordon, R. D., & Irwin, D. E. (2000). The role of physical and conceptual properties in preserving object continuity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 136–150.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics.* New York: Wiley.

Growney, R., & Weisstein, N. (1972). Spatial charcteristics of metacontrast. *Journal of the Optical Society of America, 62,* 690–696.

Harris, L. R., Duke, P. A., & Kopinska, A. (2006). Flash lag in depth. *Vision Research, 46*, 2735–2742.

He, L., Zhang, J., Zhou, T., & Chen, L. (2009). Connectedness affects dot numerosity judgment: Implications for configural processing. *Psychonomic Bulletin and Review, 16*, 509–517.

He, S., Cavanagh, P., & Intriligator, J. (1996). Attentional resolution and the locus of awareness. *Nature, 383*, 334–338.

Helsen, W., Gilis, B., & Weston, M. (2006). Errors in judging "offside" in association football: Test of the optical error versus the perceptual flash-lag hypothesis. *Journal of Sports Sciences, 24*, 521–528.

Hikosaka, O., Miyauchi, S., & Shimojo, S. (1993). Focal visual attention produces illusory temporal order and motion sensation. *Vision Research, 33*, 1219–1240.

Hsieh, P.-J., Colas, J. T., & Kanwisher, N. (2011). Pop-out without awareness: Unseen feature singletons capture attention only when top-down attention is available. *Psychological Science, 22*, 1220–1226.

Hyde, D. C., & Spelke, E. S. (2011). Spatiotemporal dynamics of processing nonsymbolic number: An event-related potential source localization study. *Human Brain Mapping, 33*, 2189–2203.

Ishii, M., Seekkuarachchi, H., Tamura, H., & Tang, Z. (2004). 3D flash lag illusion. *Vision Research, 44*, 1981–1984.

Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology, 24*, 175–219.

Kanai, R., Carlson, T. A., Verstraten, F. A. J., & Walsh, V. (2009). Perceived timing of new objects and feature changes. *Journal of Vision, 9*(7):5, 1–13, http://journalofvision.org/9/7/5/, doi:10.1167/9.7.5.

Kaufman, E. L., Lord, M. W., Reese, T. W., & Volkmann, J. (1949). The discrimination of visual number. *American Journal of Psychology, 62*, 498–525.

Khurana, B., Nijhawan, R., & Watanabe, K. (1998). *Extrapolation of stroboscopically moving objects.* Presented at the Association for Research in Vision and Ophthalmology Annual Meeting, Fort Lauderdale, Florida, USA.

Koch, C., & Tsuchiya, N. (2007). Attention and consciousness: Two distinct brain process. *Trends in Cognitive Sciences, 11*, 16–22.

Lamme, V. A. F. (2003). Why visual attention and awareness are different. *Trends in Cognitive Sciences, 7,* 12–18.

Lamme, V. A. F. (2006). Towards a true neural stance on consciousness. *Trends in Cognitive Sciences, 10,* 494–501.

Lamme, V. A. F. (2010). How neuroscience will change our view on consciousness. *Cognitive Neuroscience, 1,* 204–240.

Lamme, V. A. F., Zipser, K., & Spekreijse, H. (2002). Masking interrupts figure–ground signals in V1. *Journal of Cognitive Neuroscience, 14,* 1044–1053.

Lau, H. C., & Passingham, R. E. (2006). Relative blindsight in normal observers and the neural correlate of visual consciousness. *Proceedings of the National Academy of Sciences of the United States of America, 103,* 18763–18768.

Libet, B., Alberts, W. W., Wright, E. W., Delattre, L. D., Levin, G., & Feinstein, B. (1964). Production of threshold levels of conscious sensation by electrical stimulation of human somatosensory cortex. *Journal of Neurophysiology, 27,* 546–578.

Libet, B., Wright, E. W., Feinstein, B., & Pearl, D. K. (1979). Subjective referral of the timing for a conscious sensory experience. *Brain, 102,* 193–224.

Liu, T., Fuller, S., & Carrasco, M. (2006). Attention alters the appearance of motion coherence. *Psychonomic Bulletin and Review, 13,* 1091–1096.

MacDonald, J., & McGurk, H. (1978). Visual influences on speech perception processes. *Perception and Psychophysics, 24,* 253–257.

MacKay, D. M. (1958). Perceptual stability of a stroboscopically lit visual field containing self-luminous objects. *Nature, 181,* 507–508.

Macknik, S. L., & Livingstone, M. S. (1998). Neuronal correlates of visibility and invisibility in the primate visual system. *Nature Neuroscience, 1*, 144–149.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information.* New York: W. H. Freeman and Company.

Mitroff, S. R., & Alvarez, G. A. (2007). Space and time, not surface features, guide object persistence. *Psychonomic Bulletin and Review, 14*, 1199–1204.

Mitroff, S. R., Scholl, B. J., & Wynn, K. (2005). The relationship between object files and conscious perception. *Cognition, 96*, 67–92.

Moore, C. M., & Enns, J. T. (2004). Object updating and the flash-lag effect. *Psychological Science, 15,* 866–871.

Moore, C. M., Mordkoff, J. T., & Enns, J. T. (2007). The path of least persistence: Object status mediates visual updating. *Vision Research, 47,* 1624–1630.

Moore, C. M., Stephens, T., & Hein, E. (2010). Features, as well as space and time, guide object persistence. *Psychonomic Bulletin and Review, 17*, 731–736.

Mulckhuyse, M., Talsma, D., & Theeuwes, J. (2007). Grabbing attention without knowing: Automatic capture of attention by subliminal spatial cues. *Visual Cognition, 15*, 779–788.

Müsseler, J., van der Heijden, A. H. C., & Kerzel, D. (2004). Visual space perception and action: Introductory remarks. *Visual Cognition, 11,* 129–136.

Nijhawan, R. (1994). Motion extrapolation in catching. *Nature, 370*, 256–257.

Nijhawan, R. (1997). Visual decomposition of colour through motion extrapolation. *Nature, 386,* 66–69.

Nijhawan, R. (2001). The flash-lag phenomenon: Object motion and eye movements. *Perception, 30,* 263–282.

Nijhawan, R. (2002). Neural delays, visual motion and the flash-lag effect. *Trends in Cognitive Sciences, 6*, 387–393.

Noles, N. S., Scholl, B. J., & Mitroff, S. R. (2005). The persistence of object file representations. *Perception and Psychophysics, 67*, 324–334.

Ogmen, H., Breitmeyer, B. G., & Melvin, R. (2003). The what and where in visual masking. *Vision Research, 43,* 1337–1350.

Olivers, C. N. L., & Watson, D. G. (2008). Subitizing requires attention. *Visual Cognition, 16*, 439–462.

Ono, F., & Watanabe, K. (2011). Attention can retrospectively distort visual space. *Psychological Science, 22*, 472–477.

Palmer, S. (1999). *Vision science: Photons to phenomenology.* Cambridge, MA: MIT Press.

Palomares, M., & Egeth, H. (2010). How element visibility affects visual enumeration. *Vision Research, 50*, 2000–2007.

Pascual-Leone, A., & Walsh, V. (2001). Fast backprojections from the motion to the primary visual area necessary for visual awareness. *Science, 292*, 510–512.

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

Pizlo, Z. (2001). Perception viewed as an inverse problem. *Vision Research, 41,* 3145–3161.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32*, 3–25.

Posner, M. I., & Cohen, Y. (1984). Components of visual orienting. In H. Bouma & D. G. Bouwhuis (Eds.), *Attention and performance X: Control of language processes* (pp. 531–556). Hillside, NJ: Erlbaum.

Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. *Annual Review of Neuroscience, 13*, 25–42.

Pratt, J., & Arnott, S. R. (2008). Modulating the attentional repulsion effect. *Acta Psychologica, 127*, 137–145.

Railo, H., Koivisto, M., Revonsuo, A., & Hannula, M. M. (2008). The role of attention in subitizing. *Cognition, 107*, 82–104.

Rees, G., Kreiman, G., & Koch, C. (2002). Neural correlates of consciousness in humans. *Nature Reviews Neuroscience, 3*, 261–270.

Revkin, S. K., Piazza, M., Izard, V., Cohen, L., & Dehaene, S. (2008). Does subitizing reflect numerical estimation? *Psychological Science, 19*, 607–614.

Ross, J., & Burr, D. C. (2010). Vision senses number directly. *Journal of Vision, 10*(2):10, 1–8, http://journalofvision.org/10/2/10/, doi:10.1167/10.2.10.

Scheerer, E. (1973). Integration, interruption and processing rate in visual backward masking. *Psychologische Forschung, 36*, 71–93.

Shams, L., Kamitani, Y., & Shimojo, S. (2000). What you see is what you hear. *Nature, 408*, 788.

Shams, L., Kamitani, Y., & Shimojo, S. (2002). Visual illusion induced by sound. *Cognitive Brain Research, 14*, 147–152.

Sheth, B. R., Nijhawan, R., & Shimojo, S. (2000). Changing objects lead briefly flashed ones. *Nature Neuroscience, 3*, 489–495.

Silvanto, J., Cowey, A., Lavie, N., & Walsh, V. (2005). Striate cortex (V1) activity gates awareness of motion. *Nature Neuroscience, 8,* 143–144.

Supèr, H., Spekreijse, H., & Lamme, V. A. F. (2001). Two distinct modes of sensory processing observed in monkey primary visual cortex (V1). *Nature Neuroscience, 4,* 304–310.

Suzuki, S., & Cavanagh, P. (1997). Focused attention distorts visual space: An attentional repulsion effect. *Journal of Experimental Psychology: Human Perception and Performance, 23,* 443–463.

Talgar, C. P., Pelli, D. G., & Carrasco, M. (2004). Covert attention enhances letter identification without affecting channel tuning. *Journal of Vision, 4,* 22–31.

Tas, A. C., Dodd, M. D., & Hollingworth, A. (2012). The role of surface feature continuity in object-based inhibition of return. *Visual Cognition, 20,* 29–47.

Thompson, K. G., & Schall, J. D. (1999). The detection of visual signals by macaque frontal eye field during masking. *Nature Neuroscience, 2,* 283–288.

Treisman, A. M. (1964). Verbal cues, language, and meaning in selective attention. *American Journal of Psychology, 77,* 206–219.

Trick, L. M., & Pylyshyn, Z. W. (1993). What enumeration studies can show us about spatial attention: Evidence for limited capacity preattentive processing. *Journal of Experimental Psychology: Human Perception and Performance, 19,* 331–351.

Trick, L. M., & Pylyshyn, Z. W. (1994). Why are small and large numbers enumerated differently? A limited-capacity preattentive stage in vision. *Psychological Review, 101,* 80–102.

Turatto, M., Viscovi, M., & Valsecchi, M. (2007). Attention makes moving objects be perceived to move faster. *Vision Research, 47,* 166–178.

Vetter, P., Butterworth, B., & Bahrami, B. (2008). Modulating attentional load affects numerosity estimation: Evidence against a pre-attentive subitizing mechanism. *PLoS ONE, 3*(9): e3269, 1–6, doi:10.1371/journal.pone.0003269.

Watanabe, K. (2004). Visual grouping by motion precedes the relative localization between moving and flashed stimuli. *Journal of Experimental Psychology: Human Perception and Performance, 30,* 504–512.

Watanabe, K., Nijhawan, R., Khurna, B., & Shimojo, S. (2001). Perceptual organization of moving stimuli modulates the flash-lag effect. *Journal of Experimental Psychology: Human Perception and Performance, 27,* 879–894.

Watson, A. B., Barlow, H. B., & Robson, J. G. (1983). What does the eye see best? *Nature, 302,* 419–422.

Weisstein, N., Ozog, G., & Szoc, R. (1975). A comparison and elaboration of two models of metacontrast. *Psychological Review, 82,* 325–343.

Wertheimer, M. (1923). Untersuchungen zur Lehre von der Gestalt II [Laws of organization in perceptual forms]. *Psychologische Forschung, 4,* 301–350.

Whitney, D. (2002). The influence of visual motion on perceived position. *Trends in Cognitive Sciences, 6,* 211–216.

Whitney, D., & Murakami, I. (1998). Latency difference, not spatial extrapolation. *Nature Neuroscience, 1,* 656–657.

Whitney, D., Murakami, I., & Cavanagh, P. (2000). Illusory spatial offset of a flash relative to a moving stimulus is caused by differential latencies for moving and flashed stimuli. *Vision Research, 40,* 137–149.

Wichmann, F. A., & Hill, N. J. (2001). The psychometric function: I. Fitting, sampling, and goodness of fit. *Perception and Psychophysics, 63,* 1293–1313.

Xu, X., & Liu, C. (2008). Can subitizing survive the attentional blink? An ERP study. *Neuroscience Letters, 440*, 140–144.

Yantis, S. (2008). The neural basis of selective attention: Cortical sources and targets of attentional modulation. *Current Directions in Psychological Science, 17,* 86–90.

Yeshurun, Y., & Carrasco, M. (1998). Attention improves or impairs visual performance by enhancing spatial resolution. *Nature, 396,* 72–75.

Yi, D.-J., Turk-Browne, N. B., Flombaum, J. I., Kim, M.-S., Scholl, B. J., & Chun, M. M. (2008). Spatiotemporal object continuity in human ventral visual cortex. *Proceedings of the National Academy of Sciences of the United States of America, 105,* 8840–8845.