Non-Retinotopic Reference Frames in Human Vision: A Dynamic Journey from Visual Chaos to Clarity

A Dissertation

Presented to

The Faculty of the Department of Electrical and Computer Engineering

University of Houston

In Partial Fulfillment

of the Requirements for the Degree

Doctorate of Philosophy in Electrical Engineering

by

Babak Noory Meshkate

May 2014

Non-Retinotopic Reference Frames in Human Vision: A Dynamic Journey from Visual Chaos to Clarity

Babak Noory Meshkate

Approved:

Chairman of the Committee,
Haluk Ogmen, Professor,
Electrical and Computer Engineering

Committee Members:

Bruno G. Breitmeyer, Professor,
Psychology

Jose Luis Contreras-Vidal, Professor,
Electrical and Computer Engineering

Michael H. Herzog, Professor,
Laboratory of Psychophysics, Brain Mind Institute, École Polytechnique Fédérale de Lausanne (EPFL)

Bhavin Sheth, Associate Professor,
Electrical and Computer Engineering

Suresh K. Khator, Associate Dean,
Cullen College of Engineering

Badri Roysam, Professor, Chair,
Electrical and Computer Engineering

Acknowledgements

First and foremost I owe my deepest gratitude to my advisor, Dr. Haluk Ogmen, for his contributions of time, ideas, and funding to make this dissertation possible. I am grateful for his constructive criticism and encouragement, most particularly those contained in the emails I received from him during his vacation time. It has been an honor and a privilege to work with him, and I thank him for being such an exceptional mentor.

I wish to extend my sincere gratitude to the distinct members of my doctoral committee. I thank Dr. Breitmeyer and Dr. Sheth not only for their role on the committee, but also for all that they have taught me in their respective lectures and seminars. I thank Dr. Herzog, Dr. Contreras-Vidal, and Dr. Jansen for their feedback and encouragement.

I would also like to express my gratitude to Casey for her help and unlimited patience, the members of our lab, my friends, and all those who participated in my experiments. I could not have accomplished this work without them.

And above all, I am indebted to my parents for a lifetime of unconditional support.

_____

Babak Noory Meshkate

to

Atae & Manoochehr

"You cannot depend on your eyes when your imagination is out of focus."

— Mark Twain

Non-Retinotopic Reference Frames in Human Vision: A Dynamic Journey from Visual Chaos to Clarity

An Abstract
of a
Dissertation

Presented to
The Faculty of the Department of Electrical and Computer Engineering
University of Houston

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in Electrical Engineering

by
Babak Noory Meshkate

May 2014

Abstract

The optics of the eye maps neighboring points in the environment to neighboring retinal photoreceptors, and these neighborhood relations, known as retinotopic organization, are qualitatively preserved in early visual cortical areas. Under normal viewing conditions, due to object and observer movements in the environment, the stimuli impinging on retinotopic representations are highly dynamic and unstable. Thus, understanding ecological vision requires an understanding of how visual processes operate under these dynamic conditions. Retinotopically based theories, however, are not sufficient to explain how clarity of form is achieved in a dynamic environment. Non-retinotopic theories provide an alternative to address dynamic issues associated with purely retinotopic theories. Indeed, recent studies have indicated that many visual attributes of a stimulus are computed according to non-retinotopic reference frames. While those studies show the involvement of non-retinotopic reference frames in visual computation, the nature and spatio-temporal characteristics of these reference frames remain largely unknown. The primary goal of our research was to understand the nature and spatio-temporal properties of reference frames involved in non-retinotopic computations. Our results indicate that the effect of a dynamic non-retinotopic reference frame extends over space, creating a field within which target stimuli are localized and perceived relative to the reference. The fields of neighboring dynamic reference frames interact; static neighbors do not affect the fields of dynamic references; the non-retinotopic field effect is maximized when the target and the reference stimuli are in phase; and the field strength decreases with target-reference phase shift.

The results of our visual masking experiments indicate that while masking mechanisms operate in retinotopic domain, masking effect attenuates significantly in the presence of predictable non-retinotopic reference frames. We suggest that the reference frame revealed by our studies can be better described in terms of a "field" rather than an object. Our results also indicate that the interactions between reference frames occur only when they are in motion; suggesting that the fields generated by non-retinotopic reference frames are motion-based. In conclusion, this work reveals that the dynamic nature of our visual experience should be viewed as part of the solution, rather than a problem in ecological vision.

Table of Contents

List of Figures

Chapter 1 Introduction

Survival and procreation functions rely upon sensori-motor mechanisms across a wide range of living organisms, including human beings. Naturally, information representation in the human nervous system is mostly dictated by the physics of the corresponding sensory and motor organs. Great progress has been reported over the years regarding topographical mapping of the primary cortical areas of the human brain to the respective sensori and motor organs. Yet, understanding intermediate representations between the primary sensory and motor cortical areas in the brain remains a fundamental and challenging question in neuroscience and neuroengineering. The broad goal of our research is to characterize these intermediate representations and their underlying reference frames within the human visual system.

The organization of the early visual system is retinotopic (Engel, 1994; Gardner, Merriam, Movshon, & Heeger, 2008; Sereno, Pitzalis, & Martinez, 2001; Tootell et al., 1998; Tootell et al., 1995). Spatial neighborhood relations are qualitatively preserved as points in the visible world are projected onto retinal photoreceptors, and from the retina to early areas of the visual system. However, a retinotopic representation is neither necessary nor sufficient for perception of dynamic form (Ogmen & Herzog, 2010). Indeed recent studies have indicated that for a given dynamic stimulus, visual attributes such as form (Ogmen, Otto, & Herzog, 2006), luminance (Shimozaki, Eckstein, & Thomas, 1999), color (Nishida, Watanabe, Kuriki, & Tokimoto, 2007), size (Kawabe, 2008), and motion (Boi, Ogmen, Krummenacher, Otto, & Herzog, 2009) are computed according to non-retinotopic reference frames. We hypothesize that the effect of a reference frame extends over space, creating a “reference frame field”, within which

target stimuli are localized and perceived relative to the reference frame. While these studies show the essential involvement of non-retinotopic reference frames in dynamic visual computation, none explores in depth the nature and spatio-temporal characteristics of their associated fields.

1.1 Specific Aims

The primary goal of this research was to understand the nature and properties of non-retinotopic reference frames. The first three specific aims of our research investigated the spatio-temporal characteristics of these reference frame fields:

> ***Specific Aim 1:*** To determine the spatial and temporal properties of non-retinotopic reference frame fields in the human visual system.
>
> ***Specific Aim 2:*** To study the nature of interactions between the fields of two neighboring non-retinotopic reference frames.
>
> ***Specific Aim 3:*** To investigate the effect of attention on the strength of the field induced by non-retinotopic reference frames.

According to our theoretical framework (see section 2.5), stimulus information is transferred from retinotopic to non-retinotopic representation, where it is integrated and processed over time to produce clarity of moving form. The fourth specific aim of our research investigated the transfer of information between these representations:

> ***Specific Aim 4:*** To determine the factors controlling the transfer of information from retinotopic to non-retinotopic representation in the presence of deblurring mechanisms, such as visual masking.

We used several experimental paradigms to achieve our aims. We used Ternus-Pikler apparent motion (Petersik & Rice, 2006; Pikler, 1917; Ternus, 1926) to study the spatio-temporal properties of non-retinotopic representations and their effect in motion perception. We also made use of visual masking (Bachmann, 1994; Breitmeyer & Ogmen, 2006) in combination with Ternus-Pikler paradigm to investigate the processes involved in the transfer of information from retinotopic to non-retinotopic space. We expect our results to have a significant impact on our understanding of the underlying mechanisms for computation of dynamic form.

1.2 Significance

Visual processes involved in the perception of dynamic form have puzzled vision researchers for years. Significant progress has been reported in other fields of vision science such as perception of static form, detection of motion, and determination of motion direction and speed, yet very little is known about perception of moving form. The retinotopic representation of a moving target is generally smeared, constantly changing, and possibly incomplete (Ogmen & Herzog, 2010). Yet the non-retinotopic representation or our percept of an object in motion is normally sharp, stable, and complete. Answering the very fundamental question of how clarity of dynamic form is achieved under natural ecological settings can lead to a very deep impact on vision science.

From a neuroscientific perspective, great progress has been reported in topographical mapping of the primary sensori and motor organs to their corresponding cortical areas. Yet, information transfer and intermediate representations between these

cortical areas of the central nervous system remain abstract. Understanding the transfer of information from retinotopic to non-retinotopic space can reveal, to some extent, the nature of the intermediate cortical representations within the human visual system.

From a neuroengineering perspective, understanding biological vision can lead to bio-inspired reverse-engineered designs. The clarity of dynamic form perception under natural ecological settings hints at utilization of a biological signal "read-out" or signal "injection" for dynamic noise cancelation. Understanding such biological mechanisms has significant potential implications for improving the existing computational algorithms in machine vision. Understanding the intermediate representations will shed light on the visual mechanisms involved in reconstruction of space-time perceptual models from incomplete and constantly changing retinal samples of moving distal stimuli. From an engineering perspective, understanding the biological solution for reconstruction of space-time perceptual models from incomplete and smeared samples can bring about significant improvement to existing computer vision algorithms and motion picture compression techniques. Understanding spatio-temporal characteristics of non-retinotopic reference frames will further shed light on concepts of perceptual object identity and persistence, with potential implications for improving digital image classification and video indexed search algorithms.

Chapter 2 Background

At every moment in time, through the optics of the eye, a two dimensional image of the world is sampled on our retina. Under static conditions, image formation and representation in the visual system resembles that of a still camera. The nature of our visual system, however, is highly dynamic. Objects move and their respective motion imposes significant changes upon the images sampled on our retina. Furthermore, our eyes, head, and body also move and change the coordinates of objects with respect to our visual system. To make things more complicated, objects frequently move behind other objects in our visual field, and parts of the moving objects or objects behind them may be rendered invisible due to occlusion. Considering these facts, how does our visual system create and maintain a spatio-temporally continuous percept of the world from a series of incomplete and constantly changing retinal images? How are the physical locations and shapes of objects calculated with such accuracy, despite the highly transient and dynamic nature of neural stimulation? These fundamental questions have puzzled vision researchers for years. It is widely accepted that the visual system creates high-level models of objects in sight (Kahneman, Treisman, & Gibbs, 1992). Rather than analyzing every retinal image sample, these models are created, updated, and deleted upon significant change observed across several retinal samples. Significance of visual reference frames in object recognition and modeling has been studied extensively (Biederman, 1987; Biederman & Gerhardstein, 1993; Dunker, 1929; Johansson, 1973, 1975; Rock, Auster, Schiffman, & Wheeler, 1980; Rock & Divita, 1987; Tarr, 1995; Tarr & Pinker, 1989, 1990). Perceptual reference frames are analogous to geometric coordinate systems, in the sense that they can map a spatial object into different

representations. Such perceptual reference frames can be categorized as ego-centric (viewer-centered) or exo-centric (non-viewer-centered). Retinotopic reference frames are examples of ego-centric reference frames, while object-based, space-based, or motion-segmentation-based non-retinotopic reference frames are examples of exo-centric reference frames.

2.1 Retinotopic vs. Non-Retinotopic Representation

Most theoretical accounts of visual perception have been constructed around the concept of *retinotopic organization* of human visual system. *Retinotopy* refers to qualitative preservation of spatial relations between points in the visible world, as the image of distal visual stimuli is orderly projected from the retina to the early areas of the visual system such as the LGN, V1, etc. (Engel, 1994; Gardner et al., 2008; Sereno et al., 2001; Tootell et al., 1998; Tootell et al., 1995). The nature of the human visual system, however, is highly dynamic. Voluntary and involuntary eye movements for instance cause the retinal image to undergo significant changes over time. Under normal viewing conditions, we make saccadic eye movements for gaze repositioning at an average frequency of 3Hz. Average speeds of saccades are in the order of hundreds of arc-degrees per second, and the resulting retinal image motion is drastic. Nevertheless, we perceive the world as uniform, continuous, and stable. On the other hand, object motion poses another significant challenge to purely retinotopic theories of human vision. Under normal viewing conditions, a briefly presented stimulus remains visible for approximately 120 ms, after the stimulus offset (Coltheart, 1980; Haber & Standing, 1970). Due to this phenomenon, formally known as visible persistence, retinal samples of moving objects become smeared and overlapping (Figure 1). Faded and overlapping

copies of fast moving objects are expected to spread across retinal space leading to a smeared image. Yet our perception of such moving objects remains relatively sharp and clear. These examples illustrate that retinotopic representation is not sufficient for perception of dynamic form (Ogmen & Herzog, 2010). In the following sections, we will look at perception of moving form and its associated complications in more detail, and explore more examples that necessitate non-retinotopic processing of visual stimuli.

2.2 Perception of Moving Form

As mentioned in the previous section, visible persistence of a briefly presented stationary stimulus under normal viewing conditions has been shown to be 120 ms. Existence of visible persistence after presentation of stationary stimuli leads to expectation of a comet-like trailing smear following moving objects. Yet our perception of objects in motion is relatively clear and sharp (Bex, Edgar, & Smith, 1995; Hammett, 1997; Ramachandran, Rao, & Vidyasag, 1974). Purely retinotopic theories of vision fail to explain how biological visual systems produce a clear percept of moving objects from such distorted retinal samples. While retinotopic stabilization of moving objects can be achieved in smooth pursuit eye movement, the problem of motion smear and interference in the case of pursuit movement is simply transferred to other objects simultaneously present in the visual field (Bedell & Lott, 1996). We suggest that the visual system resolves these problems at two distinct levels: i) reduction of spatial extent of motion smear is performed in *retinotopic space*; ii) computation of clear form is performed in *non-retinotopic space* (Ogmen, 2007).

2.3 Retinotopic Control of Spatial Extent of Motion Smear – Motion Deblurring

Perceived extent of motion smear produced by moving dots as a function of exposure duration (Burr, 1980; Dilollo & Hogben, 1985) indicates that for exposure durations under approximately 40 ms, perceived smear increases with exposure duration as expected. However, for longer exposure durations (> 40 ms), the extent of perceived smear is significantly less than that predicted from the persistence of static stimuli. This reduction of perceived smear for moving objects in the human visual system is known as *motion deblurring* (Burr & Morgan, 1997). Despite the reports of motion deblurring, motion smear has been observed for isolated targets in both real (Bidwell, 1899; McDougall, 1904) and apparent motion (Castet, 1994; Dilollo & Hogben, 1985; Dixon & Hammond, 1972; Farrell, 1984; Farrell, Pavel, & Sperling, 1990). Further research shows that motion deblurring in a target object is a by-product of presence of other spatio-temporally proximal stimuli (Chen, Bedell, & Ogmen, 1995). Several motion estimation/compensation models have been proposed as mechanisms involved in the process of motion deblurring (Anderson & Vanessen, 1987; Burr, 1980; Paakkonen & Morgan, 1994). Nevertheless, existence of motion smear in the absence of neighboring proximal stimuli strongly contradicts existence of motion estimation/compensation mechanisms within human visual system. Inhibitory mechanisms have also been suggested for motion deblurring (Castet, 1994; Dilollo & Hogben, 1985; Dixon & Hammond, 1972; Francis, Grossberg, & Mingolla, 1994; McDougall, 1904; Ogmen, 1993).

Dependence of motion deblurring on stimulus timing and luminance bears striking resemblance to that of metacontrast masking, a visual phenomenon in which visibility of a target stimulus is reduced by a spatially non-overlapping and temporally following mask stimulus (Bachmann, 1994; Breitmeyer & Ogmen, 2006). Duration of visible persistence in apparent motion (Castet, 1994; Dilollo & Hogben, 1985; Farrell, 1984) and target visibility in metacontrast masking (Alpern, 1952; Lefton, 1973) both decrease as the spatial separation between successively presented targets is reduced. Occurrence of motion deblurring for exposure durations of longer than 40 ms also corresponds with reports of optimal metacontrast masking when the mask follows the target by 40-100 ms (Breitmeyer & Ogmen, 2006). Finally, in agreement with stronger metacontrast masking in the periphery (Alpern, 1952; Stewart & Purcell, 1974), motion deblurring is stronger in the periphery than in the fovea. The REtino-COrtical Dynamics (RECOD) model (Ogmen, 1993) has been applied to both metacontrast and motion deblurring paradigms. In this model the inhibition of sustained activities by transient activities ("transient-on-sustained inhibition") has been suggested as the main inhibitory process in metacontrast. Simulation results for a wide-range of motion deblurring and metacontrast data provided strong evidence that transient-on-sustained inhibition is the main mechanism for motion deblurring (Breitmeyer & Ogmen, 2006; Purushothaman, Ogmen, Chen, & Bedell, 1998). Furthermore, clinical evidence also supports the prediction of transient-on-sustained inhibition in motion deblurring (Tassinari, Marzi, Lee, Di Lollo, & Campara, 1999). The findings on retinotopic control of spatial extent of motion smear can thus be summarized by the following statements:

I) Extensive motion blur is perceived if moving targets are isolated on a uniform background.

II) The spatial extent of perceived motion blur is reduced in the presence of spatio-temporally proximal stimuli (motion deblurring).

III) The phenomenon of motion deblurring cannot be explained by motion compensation mechanisms.

IV) Transient-on-sustained inhibition in metacontrast mechanisms can account for motion deblurring.

2.4 Non-retinotopic Processing of Dynamic Form

The aforementioned metacontrast mechanism of transient-on-sustained inhibition can shorten motion streaks and reduce the amount of blur. Yet, such deblurred moving objects would still lack clarity of form. This problem is evident in photographs of moving objects taken at low shutter speeds, a phenomenon referred to as the problem of "*moving ghosts"* (Ogmen, 2007). This problem is directly proportional to the speed of moving targets. Fast moving targets expose an area of the film or a pixel of the digital sensor in the camera very briefly, failing to provide sufficient exposure for capturing details.

Figure 1: The Moving Ghost problem depicted in a photograph taken at low shutter speeds, resembling the visible persistence of the visual system. Under normal viewing conditions, a briefly presented stimulus remains visible for approximately 120 ms, after the stimulus offset. Due to this phenomenon, formally known as visible persistence, retinal samples of moving objects become smeared and overlapping. Insufficient stimulation of retinotopic receptive fields should leave a ghost like image of the fast moving objects, similar to the truck depicted in the picture.

As the speed of the moving target approaches zero degrees per second (static target), the exposure time is long enough to capture the details of the target. Similarly, temporal exposure of fast moving targets stimulates retinotopic neurons briefly and insufficiently. Incompletely processed form information will spread across the retinotopic space, leading to generation of a ghost-like copy of the moving target.

Figure 2: Under dynamic conditions, retinotopically localized receptive fields are briefly stimulated, resulting in the loss of feature information due to insufficient exposure. As such, under fixation conditions, faded and overlapping samples of fast moving objects spread across the retinotopic space creating the "moving ghost" problem.

How do biological visual systems overcome the problem of moving ghosts? We hypothesize that information about the form of objects in motion is conveyed to a non-retinotopic space, where it can accrue over time to allow neural processing to synthesize shape information. In fact, recent studies have indicated that visual attributes of a stimulus such as form (Ogmen et al., 2006), luminance (Shimozaki et al., 1999), color (Nishida et al., 2007), size (Kawabe, 2008), and motion (Boi et al., 2009) are computed according to non-retinotopic reference frames. A non-retinotopic reference frame moves along with the target object, allowing feature information to be collected and integrated across time, resulting in a clear and sharp percept of the moving target.

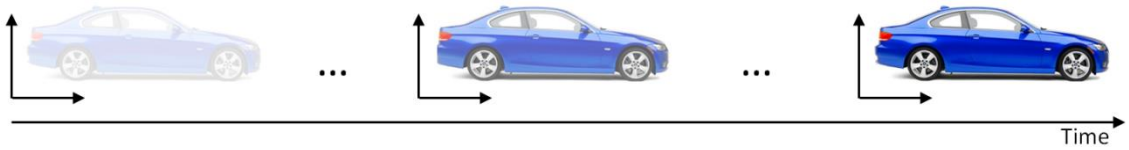


Figure 3: A non-retinotopic reference frame moves along with the target object, allowing feature information to be collected and integrated across time, resulting in a clear and sharp percept of the moving target.

2.5 Visual Reference Frames and Non-Retinotopic Manifolds

Detailed visual analysis of features on a moving object can be realized if the moving object is stabilized at a locus where feature integration occurs (e.g. the shifter-circuit model (Anderson & Van Essen, 1987). Pursuit eye movement or a single global motion compensation mechanism can stabilize the moving target, but fail to address the issue in its entirety, since the environment contains multiple objects moving with different velocities. We suggest that locally common motion vectors are computed and utilized as a reference frame to minimize motion variations within a local neighborhood. Retinotopic mapping of the visual system forms the heuristics behind this local approach (complications of occlusion will be discussed later). In addition to their local motion vectors, all parts of an object share a common motion vector. For instance, all joints in a walking human body share a common motion vector, regardless of their local motion (Johansson, 1973). The common motion vectors in local retinotopic neighborhoods can be used to partially stabilize moving objects. Role of common motion in image grouping and segmentation, and the attendant relativity of motion have extensive empirical support (Cutting & Proffitt, 1982; Johansson, 1973, 1975; Kalveram & Ritter, 1979; Restle, 1979; Wallach, 1959; Wallach, Nitzberg, & Becklen, 1985). Yet the exact rules of how reference frames act upon other stimuli are poorly understood. In this study, we investigated the nature of motion-based reference frames. Qualitative preservation of neighborhood relations simplifies the geometry and analysis of retinotopic representations. On the other hand, non-retinotopic representation is complex, and its geometry and underlying mechanisms remain generally unknown. In our visual masking experiments, we investigated the transformation of retinotopic representations to non-

retinotopic ones. Our research is based on the hypothesis that retinotopic relations for objects in motion are preserved in non-retinotopic space by means of perceptual manifolds. These manifolds are built based on extraction of common motion vectors, which serve as local reference frames for the respective manifolds. We hypothesize that "stabilization" of moving stimuli with respect to neuron receptive fields for detailed feature analysis is achieved in two steps: i) computation of common motion vectors which serve as reference frames for the non-retinotopic manifold; ii) computation of residual motion vectors with respect to the local reference. Figure 4 illustrates our hypothesized use of common motion vectors and local reference frames for transfer of visual information from retinotopic space to non-retinotopic representations. At the bottom of the figure, in the retinotopic space a group of dots move rightwards (highlighted in red) while a group of dots move upwards (highlighted in orange). Based on differences in motion vectors, the two local neighborhoods are mapped into two different non-retinotopic representations; for clarity, the figure shows only the non-retinotopic representation for rightward moving dots. A local neighborhood common motion vector (dashed green vector) is determined and utilized as the reference frame for all targets within that neighborhood. Accordingly, all motion vectors are decomposed into a sum of the reference motion and a residual motion vector. The stimulus in the local neighborhood is mapped onto a manifold (for depiction purposes, a sphere is used), i.e. a geometric structure that preserves local neighborhood relations, while the surface can be stretched and deformed.

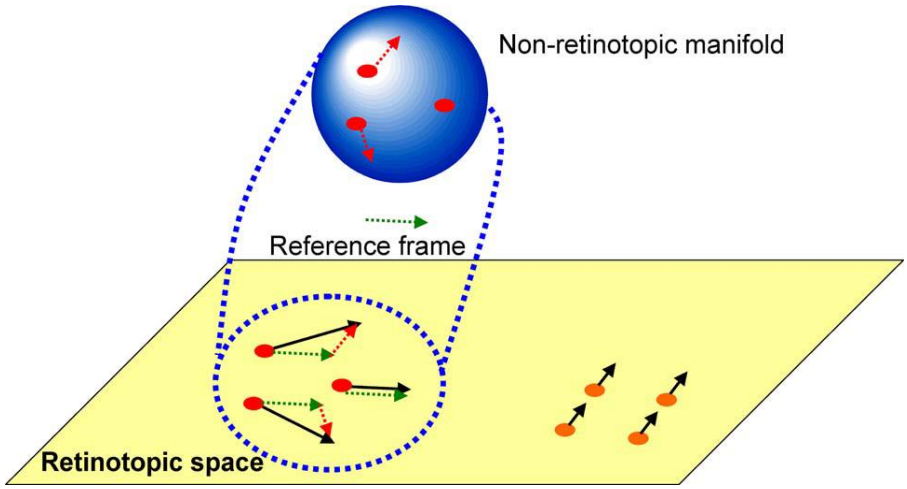


Figure 4: Retinotopic to Non-Retinotopic Transfer: Local common motion vectors in retinotopic space are extracted and used as a reference frame for transfer of visual information from retinotopic to non-retinotopic space. The residual motion vectors, or relative motion with respect to the reference frame, are then applied to the manifold so as to deform it to induce transformation that the shape undergoes during motion.

The residual motion vectors, or relative motion with respect to the reference frame, are then applied to the manifold so as to deform it to induce transformation that the shape undergoes during motion. The higher the speed of a stimulus, the briefer is the stimulation it exerts on a retinotopically localized neuron. In turn, the briefer the stimulation, the weaker the retinotopic signal. The weaker the retinotopic signal mapped from retinotopic to non-retinotopic space, the longer is the integration needed to achieve a robust signal-to-noise ratio to compute form.

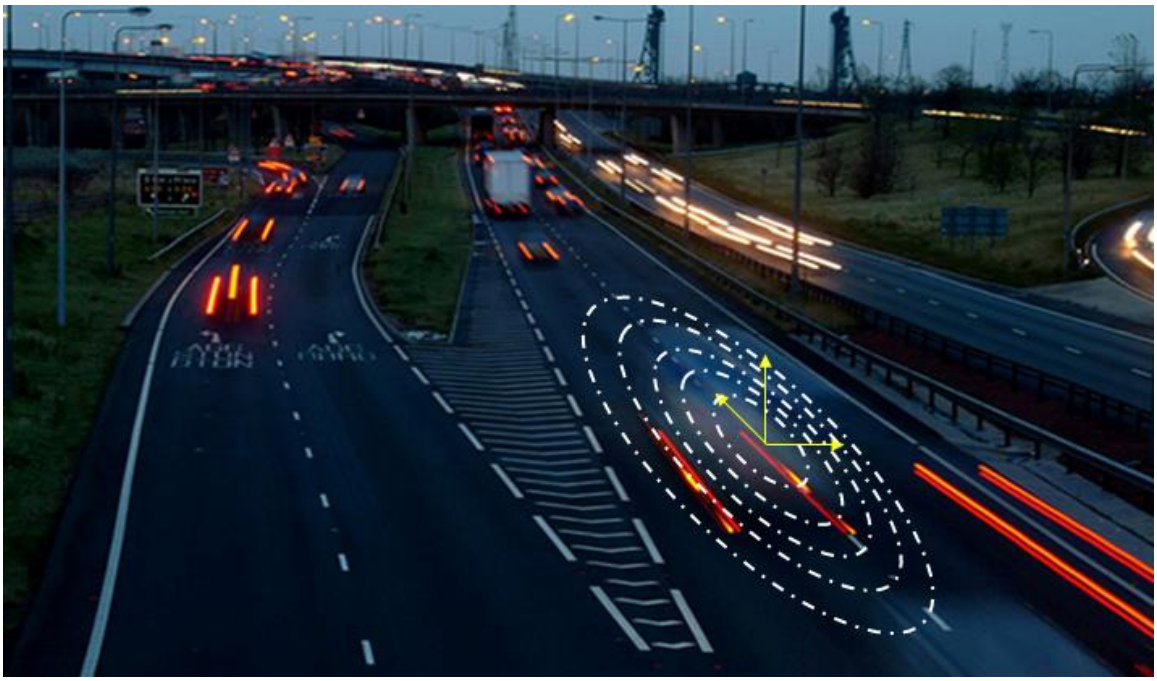


Figure 5: The effect of the non-retinotopic reference frame is hypothesized to extend in time and space as a field, within which targets can be stabilized and perceived relative to the reference.

2.6 Experimental Paradigms for Exploring Retinotopic vs. Non-Retinotopic Processing

Several paradigms have been suggested for exploring non-retinotopic visual processing. In this section, we will take a closer look at three of such methodologies and briefly discuss their advantages and disadvantages.

2.6.1 Saccadic Stimulus Presentation Paradigm

Saccadic Stimulus Presentation Paradigm (SSPP) has been the classical experimental technique used to pit retinotopic against non-retinotopic processes (Davidson, Fox, & Dick, 1973; Irwin, 1991; Knapen, Rolfs, & Cavanagh, 2009; McRae, Butler, & Popiel, 1987; Melcher & Colby, 2008; Melcher & Morrone, 2003). In a typical

SSPP experiment, two spatially overlapping but temporally separated stimuli are presented to the subject immediately before and after a saccade. Since the respective stimulated retinal regions for the two stimuli are distinct due to the saccadic eye movement, retinotopic processing theories predict no interaction between the respective percepts. Spatiotopic processing theories, on the other hand, predict significant interaction as both stimuli share the same region in space.

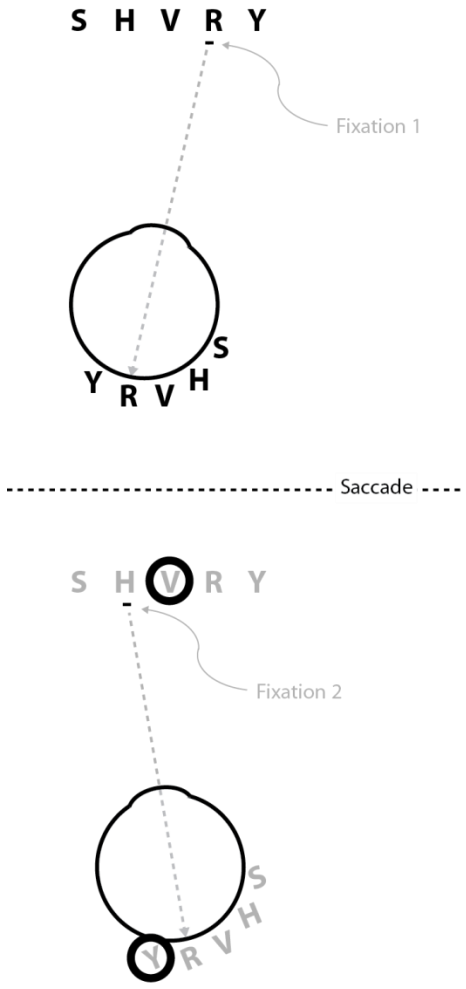


Figure 6: Saccadic Stimulus Presentation Paradigm (SSPP) used by Davidson et al. (1973). The observer makes a saccade from the first to the second fixation. Target stimuli, consisting of five letters are presented briefly just before the initiation of the saccade. A ring mask is presented after the saccade. The light gray target

letters at the bottom show the relative position of the mask with respect to the target letters. In the actual stimulus, letters were only presented before the saccade. As one can see from the figure, the ring mask surrounds letter V according to spatiotopic coordinates and letter Y according to retinotopic coordinates. The non-overlapping ring mask corresponds to the metacontrast condition. The experiments also had an overlapping pattern to examine masking by structure.

SSPP provides a powerful method for exploring non-retinotopic processing across saccades. However, this paradigm involves eye-movement related processes, such as saccadic suppression and efference copy, and cannot be employed to study non-retinotopic reference frames independent of eye movements.

2.6.2 Object-Specific Preview Paradigm

In a series of experiments, Kahneman et al. introduced the concept of Object-Specific Preview Advantage (OSPA), and laid the foundation for what became known as the Object File Theory (Kahneman et al., 1992). Kahneman et al. hypothesize that, by processing of a stationary scene, an Object File (OF) is created for every object in view. An object file is defined as a temporary episodic representation, within which successive states of an object are linked and integrated. The object files are assumed to be addressed by location at a particular time. They collect and store information about the specific object, and remain open as long as the object is in view or shortly thereafter. In order to provide perceptual continuity through dynamic change, the changed visual input must be matched against the information stored in memory prior to change. Object file theory assumes three necessary operations to achieve this goal. The *correspondence* operation is hypothesized to determine whether the object in the terminal display is “new” or a viewed object, previously perceived at a different location. A *reviewing* process is hence

necessary to retrieve, from its corresponding object file, the characteristics of the initial object now no longer in view. An *impletion* process is then assumed to use current and reviewed information for producing a percept of change or motion to link those views.

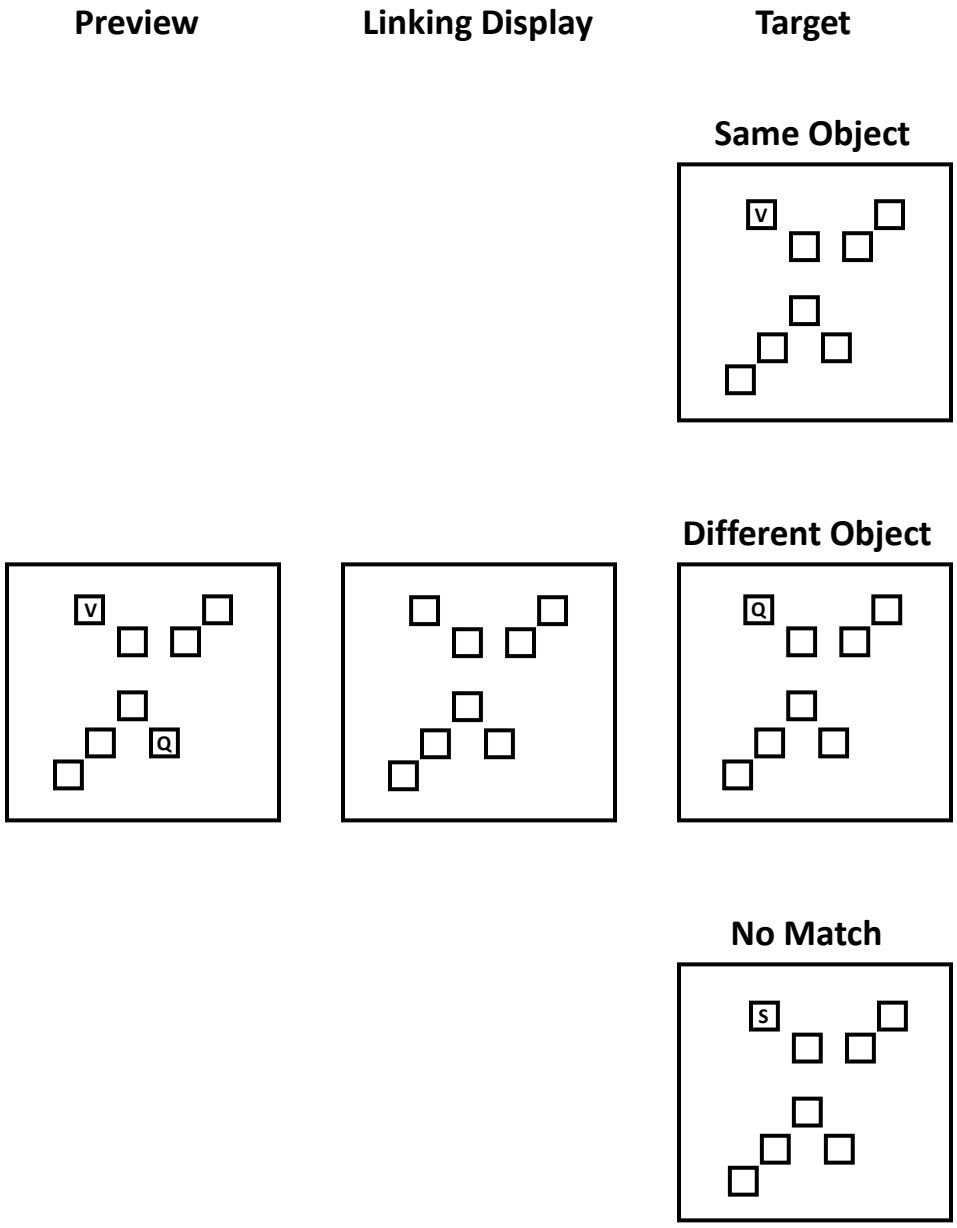


Figure 7: Stimulus used to formulate the object-file theory (Kahneman et al., 1992). Facilitation of letter naming latency is defined as the difference between SO and DO conditions and termed the Object Specific Preview Advantage (OSPA).

Kahneman et al. used facilitation in letter naming latency to investigate whether or not newly appearing objects are matched to past appearances by means of the information stored in an open object file (Kahneman et al., 1992). Their general experimental paradigm (Figure 7) consisted of two successive displays (preview/target). The preview display contained two or more letters, while the target display contained only one. The task of the observer was to name the target letter as quickly as possible. Note that the static square frames which remained visible throughout the experiment were used to define visual objects. Three experimental conditions were examined in their experiments. In the Same Object (SO) condition, the target letter matches the preview letter seen as belonging to the same object. In the Different Object (DO) condition, the target letter matches a preview letter seen as belonging to a different object. And in the No Match (NM) condition the target letter matches none of the preview letters. Kahneman et al. found the latencies to be quite similar in conditions DO and NM, and significantly faster in condition SO. They termed the facilitation of letter naming latency found as the difference between SO and DO conditions the Object Specific Preview Advantage (OSPA). The central result of their experiments was the finding of a substantial object-specific preview advantage. The standard account of visual priming predicts no difference in response latency, regardless of which square frame the matching prime appears in. However, Kahneman et al. found object-specific preview advantage to be substantially higher than the non-specific preview benefit, defined as the difference between DO and NM conditions. Furthermore, their results showed that in case of longer preview duration (250ms), the preview effect becomes almost entirely object specific.

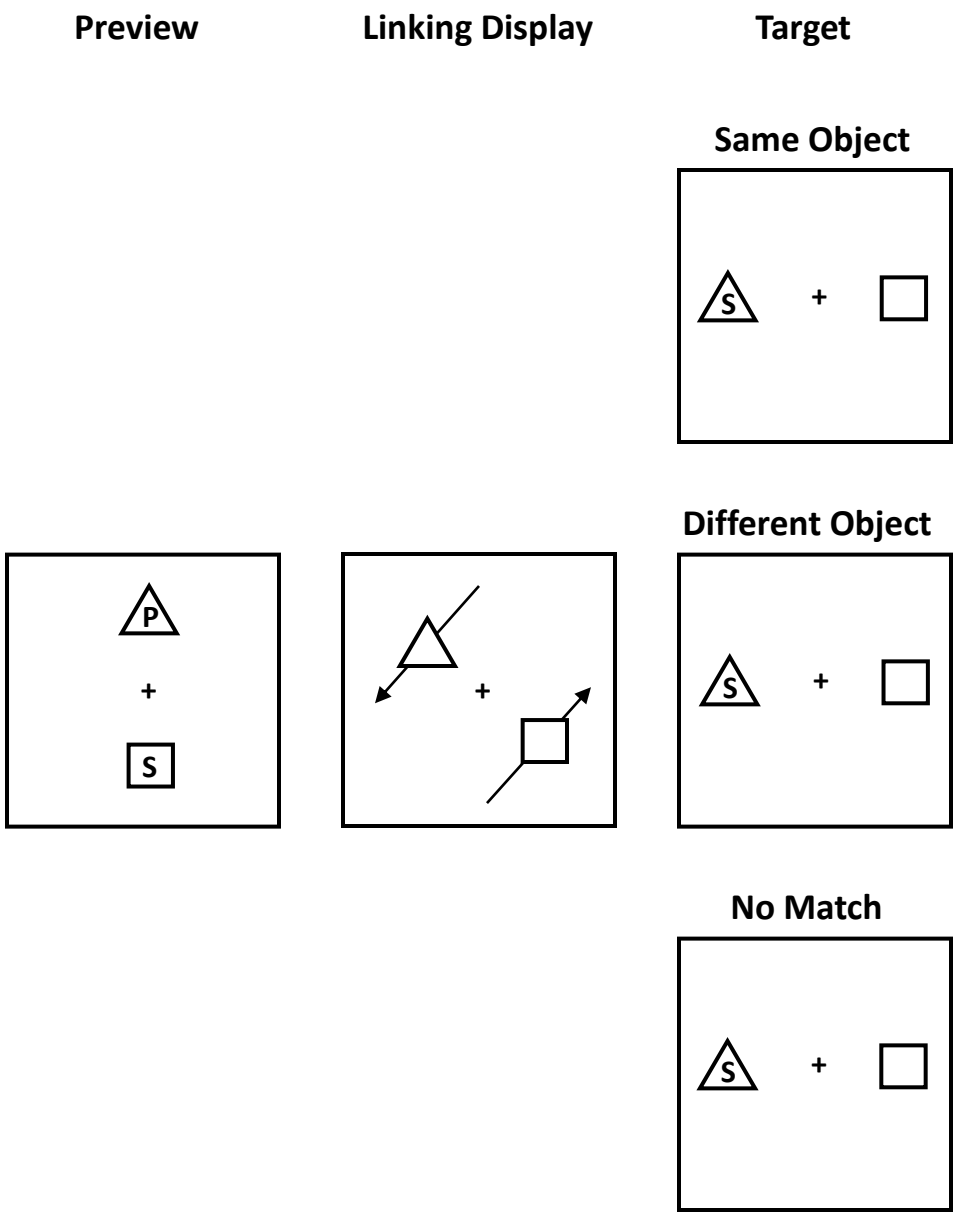


Figure 8: Dynamic Stimulus used by Kahneman et al. to explore object based priming (Kahneman et al., 1992). Motion of the square and triangular objects serve as non-retinotopic links leading to the OSPA effect.

In order to rule out the possibility of location specific node priming, Kahneman et al. conducted a set of experiments using a dynamic version of their paradigm as depicted in Figure 8. Using objects in motion, they dissociate possibility of the object-specific effect arising from persistence of information tied to particular spatial locations. In this

paradigm, the motion of the empty frames presented in the linking display maintains the object identity of the frames across time and space. Significant object-specific benefit was found, indicating that the preview benefit is not location specific as target and preview letters never shared the same spatial location in this paradigm. In a recent study, Lin and He (2012) investigated the retinotopy of masking by using a modified version of object-specific reviewing paradigm. A rectangular object (frame) was presented for a preview period of 200 ms. The target was presented during the last 10 ms of this preview period in one of the two sides of the rectangle. This rectangular frame was then shifted to a new location and displayed for another 200ms. The mask stimuli were presented during the first 30 ms of the shifted frame. One side of the frame contained a weak mask and the other side contained a strong mask. Neither mask occupied the same retinotopic location as the target, but one of the masks occupied the same rectangle-relative position as the target (i.e., the same side). Observers performed worse when the strong mask occupied the same relative position as the target. Lin and He interpreted this finding as evidence for non-retinotopic frame-centered backward masking. While this interpretation is plausible, it is difficult to make inferences about masking without observing the complete masking functions and comparing directly retinotopic, non-retinotopic, and baseline conditions. At the single short SOA of 10 ms (corresponding to ISI = 0 ms) used in the experiment, it is difficult to assess whether the difference in performance across the two mask types is due to masking *per se* or other factors. A baseline no-mask measure, as well as multiple SOA values can reveal the full typical type-A and type-B masking functions, and compare directly retinotopic and non-retinotopic masking conditions.

2.6.3 Ternus-Pikler Paradigm

The Ternus-Pikler display is an apparent motion stimulus, introduced by Gestalt psychologists about a century ago and employed extensively since then to study the spatio-temporal aspects of human vision (Petersik & Rice, 2006; Pikler, 1917; Ternus, 1926). The basic Ternus-Pikler paradigm (Figure 9-A) consists of two display frames separated by a blank frame presented for the duration of the Inter Stimulus Interval (ISI).

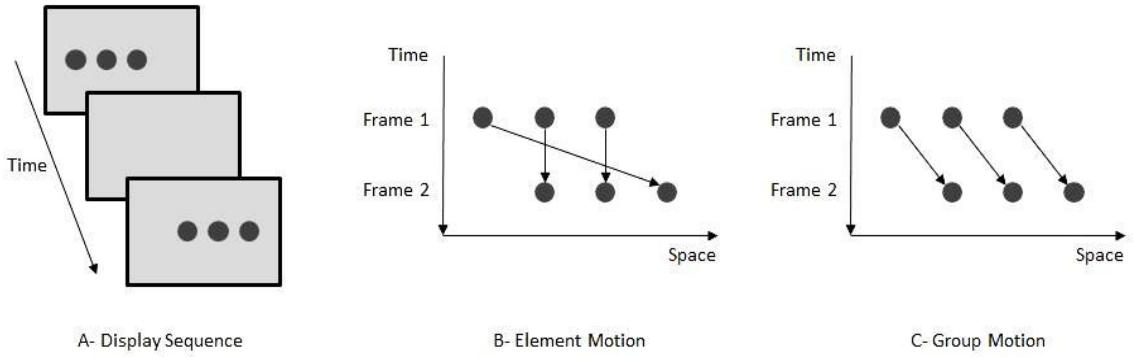


Figure 9: Ternus-Pikler display sequence and space-time diagrams for the associated motion percepts. (A) Ternus-Pikler Display: two display frames are separated by a blank interval called Inter Stimulus Interval (ISI). The two display frames are identical, except that all elements in Frame 2 are shifted by one inter element distance, with respect to the elements in the first frame. (B) Element Motion: For short ISIs (e.g., 0 ms) observers perceive the leftmost element in Frame 1 to be moving to the position of the rightmost element in Frame 2. In this case, no motion is perceived for the other two elements. (C) Group Motion: For long ISIs (e.g., 100 ms) observers perceive all elements to be moving as a group.

The two display frames are identical except that all elements in Frame 2 are shifted by one inter element distance with respect to the elements in the first frame. Depending upon the ISI duration, two different types of motion are perceived (Pantle & Picciano, 1976). For short ISIs (e.g., 0 ms), observers perceive *Element Motion*, in which

the leftmost element in Frame 1 is perceived to be moving to the position of the rightmost element in Frame 2 (Figure 9-B). In this case, no motion is perceived for the other two elements. For long ISIs (e.g., 100 ms), observers perceive *Group Motion*, in which all elements are perceived to be moving together as a group (Figure 9-C).

Several explanations have been proposed for the Ternus-Pikler phenomenon over the years with no clear consensus on the nature of its underlying mechanisms (Petersik & Rice, 2006). In his original work, Joseph Ternus (Ternus, 1926) credited Pikler (Pikler, 1917) for discovery of the Ternus-Pikler effect, and highlighted "phenomenal identity" and the relative "role" of elements in forming a Gestalt as the main explanation for the group motion percept. Kolers (Kolers) rejected the idea that "phenomenal identity" or form processing precedes motion processing, as form is a relatively weak factor in apparent motion. Pantle and Picciano (Pantle & Picciano, 1976) and Petersik and Pantle (Petersik & Pantle, 1979) suggested different underlying motion processing systems for each motion percept. Elimination of element motion, but not group motion, under dichoptic viewing conditions was taken as evidence for the dual process hypothesis. A low-level ε-process was assumed to be responsible for producing element motion, while a higher level γ-process would be responsible for group motion. The ε- and γ-processes were further identified with Braddick's short- and long-range apparent motion processes respectively (Braddick, 1973, 1974). Braddick and Adlard (Braddick & Adlard, 1978), however, noted that short-range process is insufficient to explain the element motion in Ternus-Pikler display. For instance, the motion of the outer element would require involvement of the long-range mechanisms. Cavanagh and Mather further contested the existence of short-range and long-range motion processing mechanisms, and proposed

that a single motion system can explain earlier findings if proper distinction between first-order and second-order stimulus is made (Cavanagh & Mather, 1989). Several experiments, however, showed that both element and group motion percepts can be obtained using purely second order stimului (Patterson, Hart, & Nowak, 1991; Petersik, Hicks, & Pantle, 1978). As such, neither the short-range/long-range motion processing account, nor the first-order/second-order stimulus distinction provided sufficient explanation for the Ternus-Pikler effect. Based on the concept of response persistence, a plausible alternative hypothesis was developed at the University of Houston by Bruno Breitmeyer and colleagues (Breitmeyer & Ritter, 1986a, 1986b; Ritter & Breitmeyer, 1989). Manipulation of parameters in ways that were known to increase pattern persistence was found to increase reports of element motion, while reduction of pattern persistence was shown to increase reports of group motion. These results supported the hypothesis that temporal integration of the pattern response to the two overlapping elements contributes to signaling immobility during element motion. Kramer and Rudd (Kramer & Rudd, 1999), however, showed that element motion can be perceived in the absence of visible persistence, and perception of group motion can be achieved in the presence of strong visible persistence. Nearly seventy years after the original explanation of the effect by Joseph Ternus, once again the principles of perceptual grouping emerged as the leading explanation for the Ternus-Pikler phenomenon. Kramer and Yantis (Kramer & Yantis, 1997), in a series of experiments designed to examine context effect on strength of element and group motion, proposed competition between spatial and temporal grouping of elements responsible for perception of group and element motion respectively. Context effect strengthening spatial grouping of elements in an array was

shown to increase perception of group motion, while proximity and similarity across time (strengthening temporal grouping) favored element motion. He and Ooi (He & Ooi, 1999) found similar results and proposed the across-frame and within-frame grouping responsible for group and element motion respectively.

Notwithstanding nearly a century of plausible research, underlying psychophysical and physiological mechanisms responsible for perception of element and group motion in the Ternus-Pikler display remain unknown to date. The paradigm, however, provides a suitable test for exploring visual processing in retinotopic vs. non-retinotopic coordinates. The detailed description of this methodology will be discussed in the following section.

Chapter 3 Spatial Properties of Non-Retinotopic Reference Frame Fields

The literature reviewed in Chapter 2 suggests that low-level encoding of moving stimuli occurs in a retinotopic space, and metacontrast masking mechanisms control the spatial extent of motion blur in this retinotopic space. On the other hand, accrual and processing of form information for moving objects occur in non-retinotopic space. According to our hypothesis, the transfer of information from the retinotopic to non-retinotopic space is guided by the establishment of local reference-frames. Indeed recent studies have indicated that for a given dynamic stimulus, visual attributes such as form (Ogmen et al., 2006), luminance (Shimozaki et al., 1999), color (Nishida et al., 2007), size (Kawabe, 2008), and motion (Boi et al., 2009) are computed according to non-retinotopic reference frames. In the experiments discussed in this section, we study the properties of the fields induced by these non-retinotopic reference frames. A reference frame exerts its effect on stimuli appearing in its spatio-temporal neighborhood. This effect is present even in the absence of physical/visible connections between the target object and the objects constituting the reference. Thus, with analogy to similar phenomena in physics, where a physical body can exert an effect over another body without physical contact (e.g., gravitational and electro-magnetic effects), we use the terms “perceptual field” and “perceptual force” to describe spatio-temporal extent and strength of the interactions between perceptual reference frames and target stimuli. According to our hypothetical framework, in a multi-reference environment, the perceptual fields of neighboring reference frames interact. In this section, we will discuss

the experiments conducted to study the properties of these fields and their interactions. The Ternus-Pikler experimental paradigm was utilized to achieve the following aims:

Specific Aim 1: To determine the spatial and temporal properties of non-retinotopic reference frame fields in human visual system.

Specific Aim 2: To study the nature of interactions between the perceptual fields of two neighboring non-retinotopic reference frames.

Specific Aim 3: To find the effect of attention on the strength of the perceptual field induced by non-retinotopic reference frames.

3.1 Experimental Background

Figure 10 shows a variant of the Ternus-Pikler paradigm adapted to study non-retinotopic motion processing (Boi et al., 2009). This Ternus-Pikler stimulus includes four display frames, each of which contains three disks. The display frames are separated by blank frames. Once again, depending upon the ISI value, two types of motion are perceived between the Ternus-Pikler disks. For long ISIs (e.g., 210 ms) observers perceive the disks to be moving as a group (Figure 10-A: Group Motion). For short ISIs (e.g., 0 ms) observers perceive the leftmost disk in the first/third frame to be moving to the position of the rightmost disk in the second/fourth frame and vice versa (Figure 10-B).

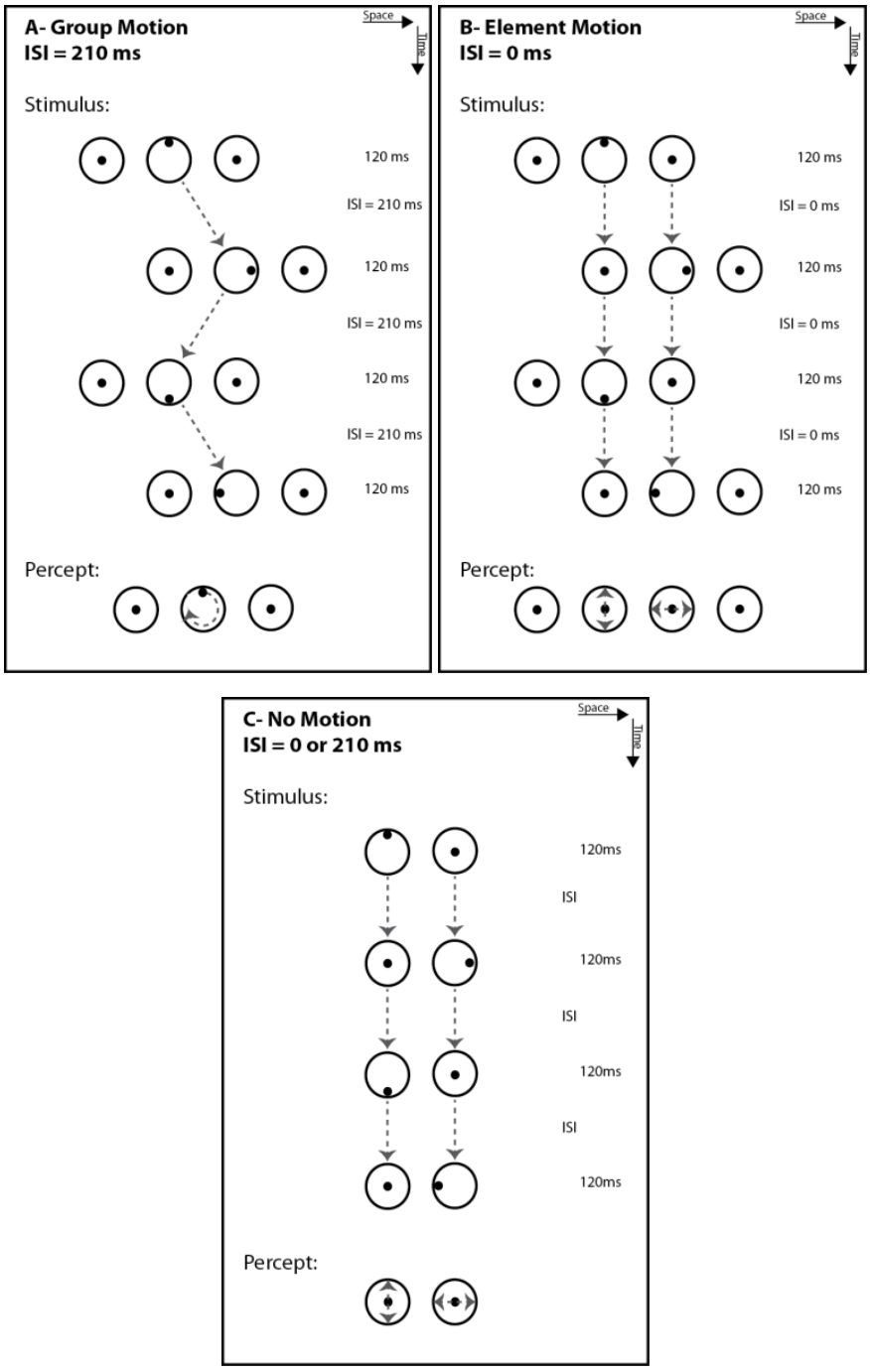


Figure 10: Stimuli and respective percepts reproduced from Boi et al.'s study (Boi et al., 2009). Ternus-Pikler space-time diagrams for three different motion conditions are depicted. Each diagram includes four display frames separated by a blank frame for the duration of ISI. (A) Group Motion: For long ISIs (e.g., 210 ms) disks are perceived to be moving as a group and the dot to be rotating. (B) Element Motion: For short ISIs (e.g., 0 ms) the leftmost disk in the first/third frame is perceived to be moving to the position of the rightmost disk in the second/fourth frame and vice versa. In this case, no motion is perceived for the

other two disks, and the dots are perceived as moving up-down and left-right inside the disks. (C) No Motion: The leftmost and the rightmost reference disks are removed, and no motion is perceived for the reference disks. The dots are perceived as moving up-down and left-right regardless of the ISI value.

In the case of element motion, no motion is perceived for the other two disks. Finally, in the no-motion control condition (Figure 10-C), removing the leftmost and the rightmost reference disks in the Ternus-Pikler display frames eliminates perception of both group and element motion, regardless of the ISI value. The percept in this case is that of two static or flickering disks. The black dots, depicted inside the Ternus-Pikler disks in Figure 10, are the probe stimuli for exploring motion perception. The solution to the general correspondence problem between these dots in consecutive frames provides a suitable benchmark test for retinotopic versus non-retinotopic motion processes. Retinotopic hypothesis predicts that the retinotopic proximity will dictate the perceived motion of the dots. Since the retinotopic proximity of subsequently presented dots in the middle disks follows the pattern shown by the arrows in Figure 10-B and -C, a purely retinotopic hypothesis predicts perception of up-down and left-right dot motion, regardless of the ISI value. Non-retinotopic hypothesis, however, predicts that the perceived dot motion depends on the perceived motion of the Ternus-Pikler disks. More specifically, the motion of the dots should be computed according to their proximity in a reference frame that moves according to the perceived motion of the Ternus-Pikler disks. In other words, the reference frame should move according to the dashed arrows in Figure 10-A and -B. When Ternus-Pikler disks are perceived to be in element motion (Figure 10-B), the non-retinotopic prediction is the same as the retinotopic prediction (perception of up-down and left-right dot motion). However, when group motion is

established between the Ternus-Pikler disks (Figure 10-A), the non-retinotopic prediction for dot motion were that of a rotation. In other words, non-retinotopic motion grouping based hypothesis predicts that group motion of Ternus-Pikler disks will serve as a non-retinotopic reference leading to the perception of dot rotation in group motion condition. The common left-right motion will serve as the reference frame and the perceived rotation of dots will depend on residual motions according to this common reference frame. Boi et al.'s results supported the predictions of non-retinotopic reference frame hypothesis.

In the present study, several variations of this paradigm were utilized to examine the nature and spatial extent of non-retinotopic reference frames involved in dynamic visual computations. Experiment 1 explores the spatial extent over which a non-retinotopic reference frame can exert its effect; Experiment 2 investigates the effect of reference size on reference frame strength, to determine if the reference frame extent scales with spatial size of the inducing elements; Experiment 3 studies the interactions between multiple reference frames; and Experiment 4 investigates the effects of endogenous attention on reference frame strength.

3.2 General Methods

All visual stimuli were generated via a Visual Stimulus Generator (VSG 2/5) card manufactured by Cambridge Research Systems. The stimuli were displayed on a 22 inch color monitor set at a resolution of 800 x 600 with a refresh rate of 100 Hz. Subject responses were collected by means of a joystick connected to the computer hosting the VSG card. The distance between the observer and the monitor was fixed at 1m, and a head/chin rest was utilized to minimize subject head motion during the experiments.

Observers were asked to maintain a stable gaze fixated at the center of the monitor and attend to the motion of the dots presented near the central disks of the Ternus-Pikler display. Eye movements was not monitored in these experiments, as previous experiments indicate that observers are able to keep a stable fixation while viewing the Ternus-Pikler displays (Boi, Ogmen, & Herzog, 2011). All experiments were conducted in a dimly lit room. Background luminance for all experiments were set at 28 cd/m^2, and the dot and disk luminance levels were fixed at 0 and 56 cd/m^2 respectively. Frame duration for all Ternus-Pikler displays was fixed at 90 ms. ISI was set at 0 and 210 ms for element and group motion conditions respectively. Several participants, majority naïve to the purpose of the study, took part in these experiments. The experiments were conducted according to a protocol approved by the University of Houston Committee for the Protection of Human Subjects. Informed consent was obtained from every participant, and practice trials were conducted to familiarize the observers with experimental procedures. The results of practice trials were not included in the data analysis.

3.3 Experiment 1: Spatial Extent of Perceptual Fields

In order to study the spatial extent of the perceptual fields induced by non-retinotopic reference frame, the distance between the dot and the Ternus-Pikler disks was varied and percent correct perceived dot rotation (clockwise/counter-clockwise) was measured. Experimental procedures and results are discussed in the following sections.

3.3.1 Experimental Methods

Stimulus design and the corresponding perceived motion for Experiment 1 are depicted in Figure 11. Center-to-center separation between the disks was fixed at 88.2'.

Disk and dot radii were also fixed at 23.79’ and 3.39’, respectively. The center-to-center distance between the black dot and the central Ternus-Pikler disc, however, was varied in the range 19.8’ to 40.74’. In addition to the element and group motion conditions, two control conditions were also included in this experiment: i) In the no-motion control condition (Figure 11-C), the outer disk/dot elements in each frame were omitted. ii) In the no-reference control condition, dots were displayed in the absence of disks. The no-reference condition is included in the study to ensure that the perception of dot rotation is in fact due to the reference Ternus-Pikler disks, and to eliminate the possibility that perception of rotation could be the result of a motion cue in the dots themselves. Observers, in a 2-AFC method, reported the perceived direction of dot rotation (clockwise or counter-clockwise). In four experimental blocks, we collected data for four Disk-Dot separations. Order of presenting the blocks within the experiment were randomized from subject to subject. Each block of the experiment included 200 trials. The trials were randomized with respect to experiment condition (motion or no-motion), ISI value (210 or 0 ms), starting position of the Ternus-Pikler motion (left or right), starting position of the target dot (top, left, right, or bottom), and the direction of dot rotation (clockwise or counter-clockwise). Each session started with the subject pressing a key on the joystick. Four display frames of 90 ms duration separated by blank frames of the appropriate ISI duration were presented in a sequence. The program then waited for the subject response, which in turn signaled the start of the next trial. Subjects were allowed to pace themselves

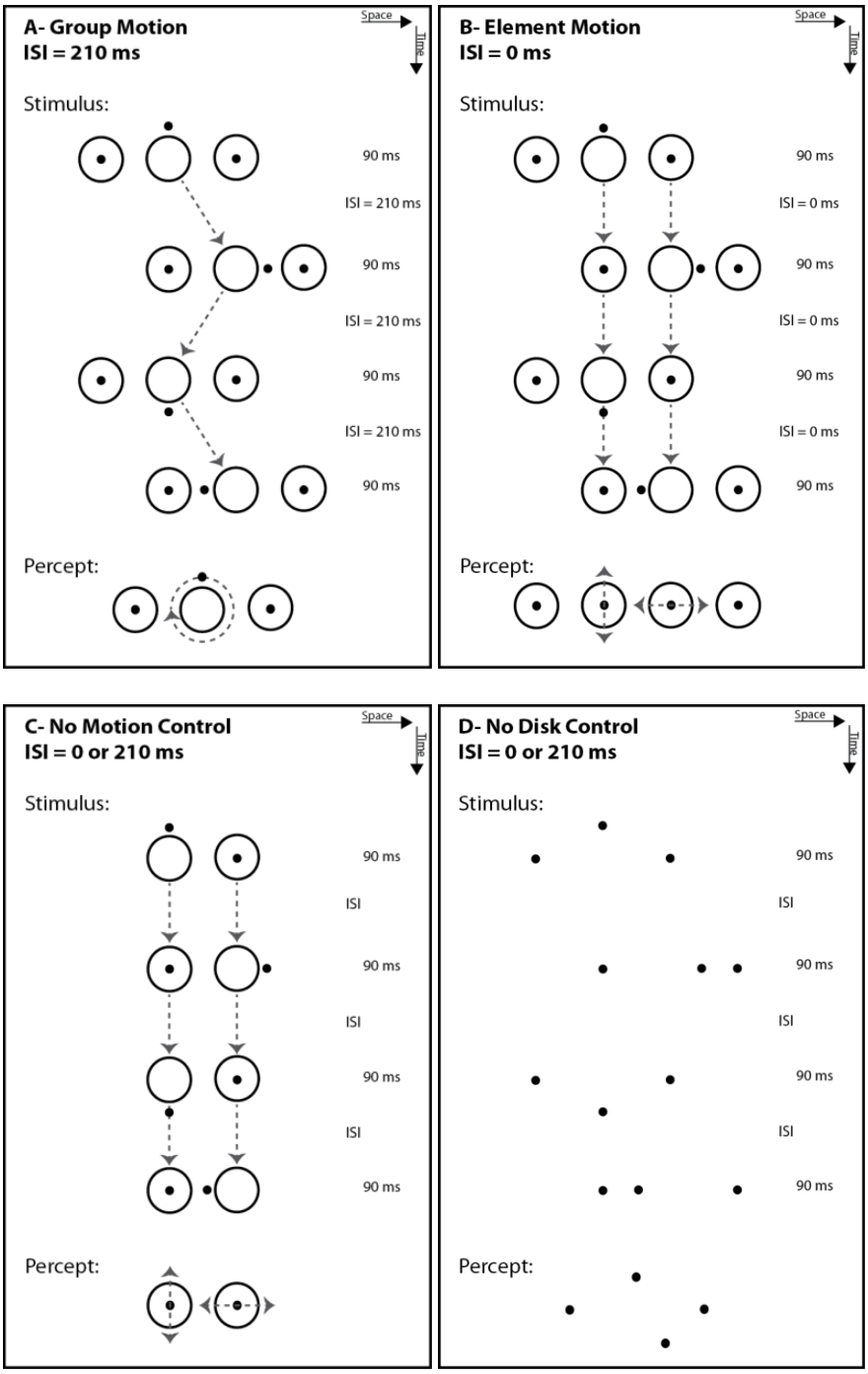


Figure 11: Stimuli for Experiments 1 and 2, and their corresponding percepts. Similar to stimuli of Figure 10, Ternus-Pikler space-time diagrams for the three different motion conditions are shown. (A) Group Motion: ISI = 210 ms. (B) Element Motion: ISI = 0 ms. (C) No Motion Control: The leftmost and the rightmost reference disks/dots removed. (D) No Reference Control: All reference disks are removed from the display, dot placement identical to (A) and (B). In Experiment 1, the center-to-center distance between the target dot and the reference Ternus-Pikler disk were varied, placing the dot inside or outside the reference disk at

different separations. In Experiment 2, the target dot was placed outside the reference disk at a fixed distance of 67.86', and the size of the reference disks were varied. Subjects were asked to report the perceived direction of rotation for the target dot.

and to take brief breaks before reporting their response. Longer rest breaks were given in between experimental blocks. The data for the no-reference control condition were collected in four independent blocks of 60 trials each.

3.3.2 Expected Results and Interpretation

Plotting performance as a function of disk-dot distance provides an estimate of the spatial extent of the reference frame perceptual field. The predicted outcomes "a", "b", and "d" shown in Figure 12 respectively correspond to strong inhibitory, medium inhibitory, and facilitatory effect of target-reference distance on perceptual field strength. Predicted outcome "c" corresponds to the case of performance independence from target-reference distance. Linear or non-linear nature of dependence for cases "a", "b", and "d" were also tested. We also analyzed weather displaying the target inside or outside the boundaries of the reference disks has an effect on performance. This has implications for "object-based" theories, as targets shown outside the boundaries of the reference disks have no connections to the disks.

We hypothesize that the perceptual strength of the field induced by the non-retinotopic reference frame decreases as the target-reference distance is increased. As such, we expect our results to be similar to prediction "a" or "b".

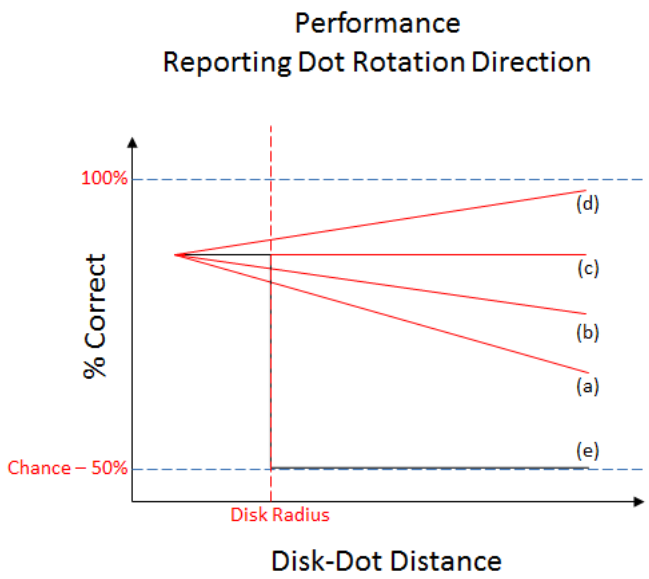


Figure 12: Hypothetical Outcomes of Experiment 1: (a) Strong inhibitory effect of distance on perceptual field strength, as performance is expected to fall near chance level at large reference-target separations. (b) Medium inhibitory effect. (c) Independence of field strength from target-reference distance. (d) Facilitatory effect of distance on perceptual field strength. (e) "Object-based" prediction, with definition of object as closed boundaries.

3.3.3 Results

Figure 13 shows performance as a function of dot-disk separation for the different experimental conditions. Two-way repeated-measures ANOVA indicates that experimental condition ($F_{3,9} = 219.7$; $p < 0.001$), but not dot-disk separation ($F_{3,9} = 3.2$; $p = 0.167$), has a significant effect on performance. When group motion condition is removed from analysis, experimental condition ceases to be significant ($F_{2,6} = 3.5$; $p = 0.157$). In fact, while the average performance was about 80% correct in the case of group motion, it is near chance for all other conditions. A paired t-test comparing performance to 50% yields: i) Element motion experiment condition ($t_{11} = 0.588$; $p = 0.567$); ii) No-motion control condition with ISI = 210 ($t_{11} = 0.580$; $p = 0.573$); iii) No-

motion control condition with ISI = 0 (t_{11} = -1.336; p = 0.208); iv) No-reference control condition ISI = 0 (t_{11} = 0.493; p = 0.625); and v) No-reference control condition ISI = 210 (t_{11} = 1.743; p = 0.091).

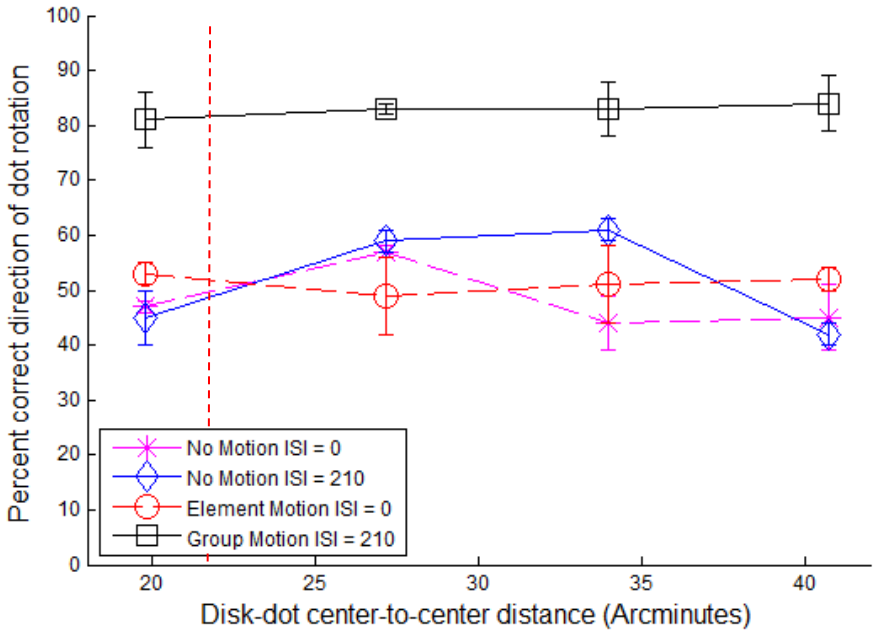


Figure 13: Results from Experiment 1: Percent correct performance in detecting direction of dot rotation, averaged across observers (N=4) and plotted against the dot-disk center-to-center distance (Arcminutes). The vertical dashed line indicates the location of the Ternus-Pikler disk boundary. The data point to its left corresponds to the case where the dot is inside the disk, while the other data points correspond to cases where the dot is outside the Ternus-Pikler disk. Subjects perform well above chance level when reference disks are perceived to be moving as a group, but near chance level for all other experimental conditions. The dot-disk separation has no significant effect on performance. Error bars correspond to ±1 SEM.

3.3.4 Discussion

The vertical dashed line depicted in Figure 13 indicates the location of the Ternus-Pikler disk boundary. The data point to its left corresponds to the case where the dot is

inside the disk, while the other data points correspond to cases where the dot is outside the Ternus-Pikler disk. In agreement with earlier findings (Boi et al., 2009), our results indicate that dot rotation is perceived only when the reference disks are perceived to be moving as a group. Furthermore, once perception of group motion is established, within the range tested in this experiment, subject performance in reporting direction of dot rotation remains independent of dot location (inside or outside the reference disk) and the disk-dot separation. This finding was contrary to our expectation of a performance drop, qualitatively proportional to the disk-dot distance. However, one must note that the Ternus-Pikler disk-disk center to center separation places an upper bound on the range of the disk-dot separation that can be tested in this paradigm. The multi-reference interaction experiments discussed in section 3.5 explores the effect of longer distances on non-retinotopic field strength. The maximum separation of 40.74' in this experiment places the dot near the half-way point between two neighboring Ternus-Pikler disks. At the maximum separation tested in this experiment, the range over which non-retinotopic reference frame effect remains constant is 12 times the radius of the dot and 1.7 times the radius of the disk. In order to investigate further the ratio of separation to inducing-element-size, we varied in the next experiment the size of the Ternus-Pikler elements.

3.4 Experiment 2: Effect of Inducing Element Size on Field Strength

In order to study dependence of non-retinotopic reference frame strength on the spatial size of inducing elements, the reference-disk size were varied and percent correct perceived dot rotation (clockwise/counter-clockwise) were measured. Experimental procedures and results are discussed in the following sections.

3.4.1 Experimental Methods

The stimulus design and methods were similar to those of Experiment 1. Dot radius and disk-dot center-to-center separation were fixed respectively at 3.39' and 67.86', while the radius size of the reference-disks were varied in the range from 6.78' to 30.63'. Observers reported the perceived direction of dot rotation in a 2-AFC method. The experimental blocks and randomization of trials were identical to those of Experiment 1, with the exception of elimination of the no-reference control condition.

3.4.2 Expected Results and Interpretation

Plotting performance as a function of disk radius provides an estimate of the effect of inducing element size on perceptual field strength (Figure 14).

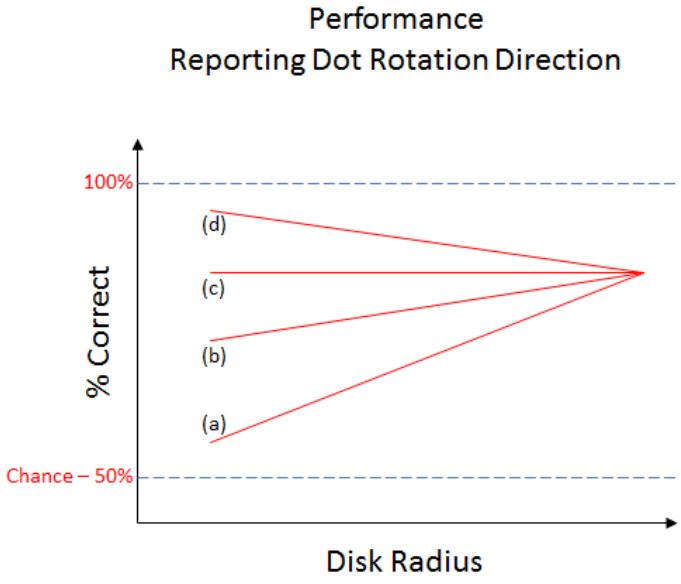


Figure 14: Hypothetical Outcomes of Experiment 2: (a) Strong facilitatory effect of inducing-element size on perceptual field strength. (b) Medium facilitatory effect. (c) Independence of field strength from inducing-element size. (d) Inhibitory effect of inducing-element size on perceptual field strength.

The predicted outcomes “a”, “b”, and “d” shown in Figure 14 respectively correspond to strong facilitatory, medium facilitatory, and inhibitory effect of inducing-element size on perceptual field strength. Predicted outcome “c” corresponds to the case of performance independence from inducing-element size. Linear or non-linear nature of dependence for cases “a”, “b”, and “d” were also tested. According to the motion-segmentation hypothesis, the motion of the inducing elements constitutes the reference frame. As such, we hypothesize that variations in disk radius have no significant effect on performance (as shown in expected results “c”).

3.4.3 Results

Figure 15 depicts the performance as a function of reference disk radius for different experimental conditions. Two-way repeated-measures ANOVA indicates that experimental condition ($F_{3,9} = 99.2$; $p < 0.001$), but not disk radius ($F_{3,9} = 1.2$; $p = 0.350$), has a significant effect on performance. When group motion condition is removed from the analysis, experimental condition ceases to be significant ($F_{2,6} = 3.0$; $p = 0.174$). Performance is above 80% correct in the case of group motion, while it is near chance for all other conditions. A paired t-test comparing performance to 50% yields: i) Element motion experiment condition ($t_{11} = -1.137$; $p = 0.279$); ii) No-motion control condition with ISI = 210 ($t_{11} = -1.355$; $p = 0.202$); and iii) No-motion control condition with ISI = 0 ($t_{11} = 0.212$; $p = 0.835$).

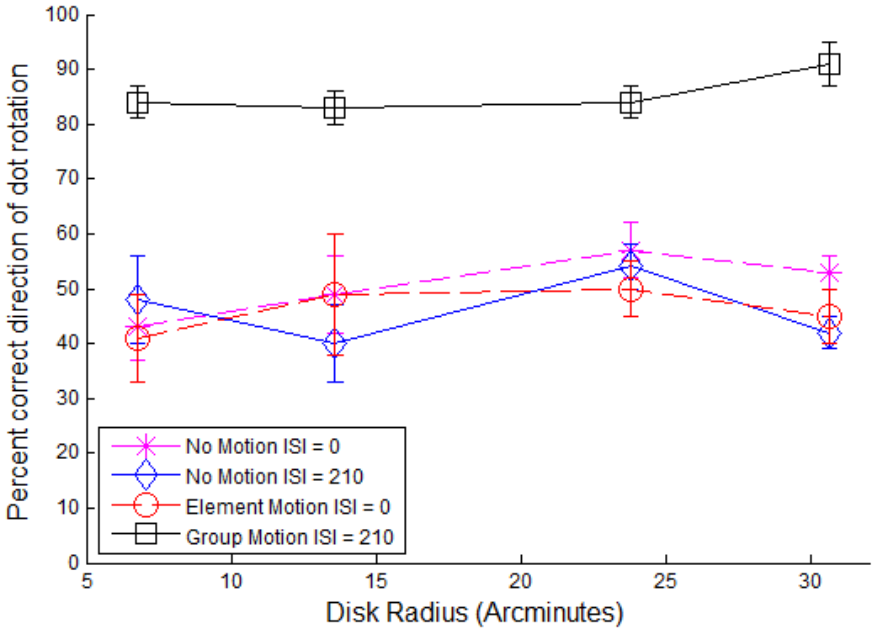


Figure 15: Results from Experiment 2: Percent correct subject performance in detecting direction of dot rotation, averaged across observers and plotted against the disk radius size (Arcminutes). The results indicate that subject performance is well above chance level when reference disks are perceived to be moving as a group, but near chance level for all other experimental conditions. The disk radius has no significant effect on performance. Error bars correspond to ±1 SEM.

3.4.4 Discussion

The results indicate that, the strength of dot rotation perception in the neighborhood of the reference frame remains independent of the reference disk size. Note that in the case of minimal disk size (6.78'), the disks are perceived to be slightly larger than the dot. Nonetheless, perception of dot rotation remains strong, so long as group motion is maintained between the reference disks. In relative terms, the constancy of the reference frame effect extends to dot-disk separation to disk-radius ratios as large as 20. In the absence of group motion, however, dot rotation perception ceases to exist. In summary, variations in spatial dimensions of dynamic objects constituting the reference frame have

no significant effect on the strength of motion induced in neighboring targets. Taken together, the results of Experiments 1 and 2 suggest that the motion of the reference elements creates a reference field that extends uniformly across a substantial spatial range. In the next experiment, we investigated how reference fields created by different reference frames interact.

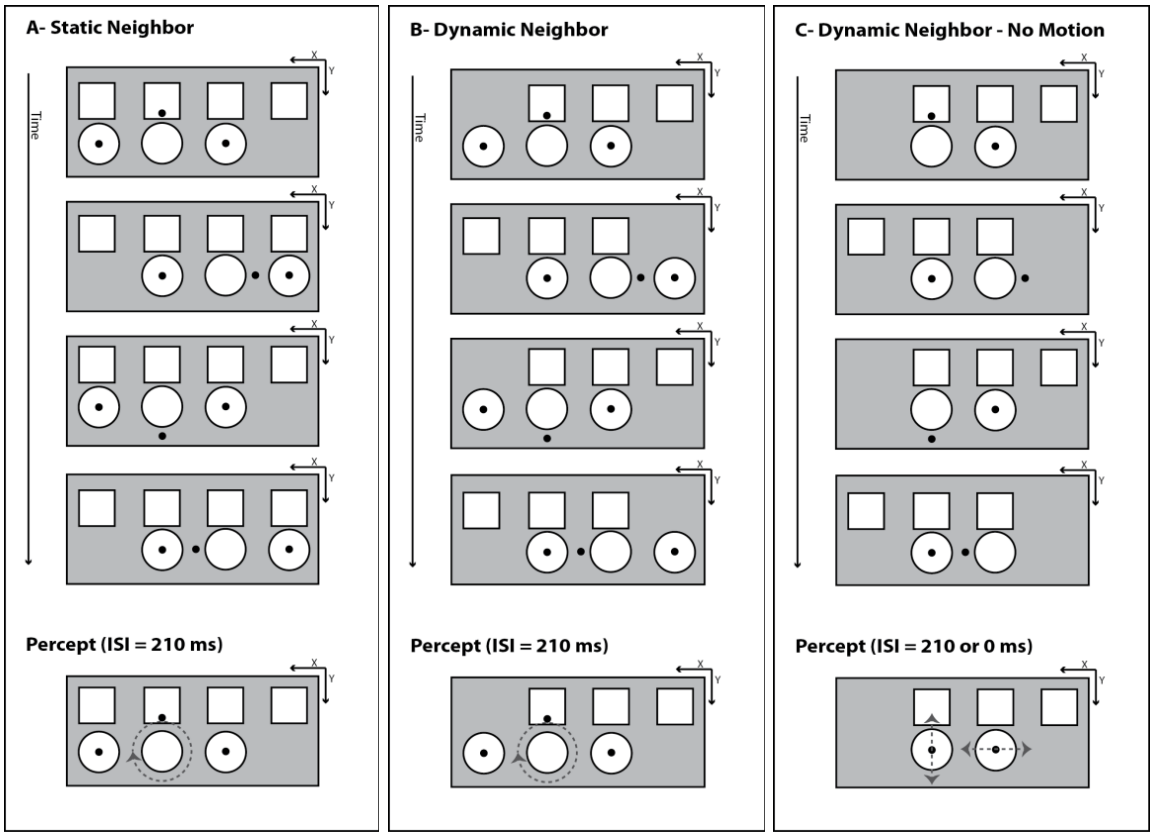


Figure 16: Stimuli used in Experiment 3 and their respective percepts. Stimuli design is similar to that of Experiment 1, with the addition of a neighboring reference frame. Different shapes (disks of 27' radius and squares of 54' sides) are chosen so that elements belonging to the two reference frames remain perceptually different from one another. Element motion and No-motion conditions are included in the experiment, but not shown here. (A) Static Neighbor: Four static squares are introduced above the Ternus-Pikler disks. The distance between the static neighboring set and the Ternus-Pikler reference is varied in the range of 67.86' to 300'. Note that at the minimal separation between the disks and squares, the target dot (4.5' radius) falls inside the boundaries of one of the neighboring squares. (B) Dynamic Neighbor: Three neighboring squares move in

a similar pattern as the Ternus-Pikler disks, but in the opposite direction. All other parameters are identical to the static neighbor condition.

3.5 Experiment 3: Interactions between Opposite-Direction Neighboring Reference Frames

In order to study the interactions between the fields of non-retinotopic reference frames in a multi-reference environment, we added a set of square objects to act as a secondary reference in our experimental paradigm (Figure 16). Experimental procedures and results are discussed in the following sections.

3.5.1 Experimental Methods

Different shapes (disks of 27' radius and squares of 54' sides) were designated so that elements belonging to the two reference frames remain perceptually different from one another. The dot radius was fixed at 4.5'. Two different conditions were examined. In the static condition, a set of four stationary squares were displayed above the Ternus-Pikler reference disks. The stationary squares appeared on the screen before the first trial and remained visible throughout the experiment. In the dynamic condition, two Ternus-Pikler stimuli, one composed of disks and one composed of squares, were displayed simultaneously. The two Ternus-Pikler displays moved in opposite direction with respect to each other, so as to create perceptual reference fields of opposite direction. The center to center distance between the squares and the disks was varied in the range from 67.86' to 300', for both static and dynamic neighborhood conditions. In a 2-AFC method, one of the authors and four naïve observers (ages 24 – 36) reported the perceived direction of dot rotation. Three of the naïve observers were chosen from the subject population of

Experiment 2. Data was collected in eight blocks (four blocks for the static and four blocks for the dynamic neighbor condition), each of which consisted of 150 trials. Since subject performance in Experiment 1 was at chance for the no-reference control condition even in the absence of an opposing neighboring field, we eliminated the no-reference control condition from Experiment 3. By analogy, no-motion control condition for ISI = 0 ms was removed, and the no-motion control condition was included for ISI = 210 ms only.

3.5.2 Expected Results and Interpretation

Plotting performance as a function of inter-reference distance provides an estimate of the interactions between perceptual fields of neighboring non-retinotopic reference frames (Figure 17). The predicted outcomes “a” and “b” correspond to inhibitory effect, “d” and “f” correspond to facilitatory effect, and “c” corresponds to no effect of inter-reference distance on performance. Linear or non-linear nature of dependence for cases “a”, “b”, “d”, and “f” were also tested. We hypothesize that the absence of motion vectors in the static neighboring reference eliminates the likelihood of any interaction between the two reference frames, and performance remains independent of inter-reference distance (prediction “c”). In the case of dynamic neighboring reference, on the other hand, we hypothesize that the inhibitory interaction between motion vectors of the two aniso-direction reference sets increases as the distance between the elements constituting the two reference frames is decreased. As such, prediction “d” is the more likely outcome.

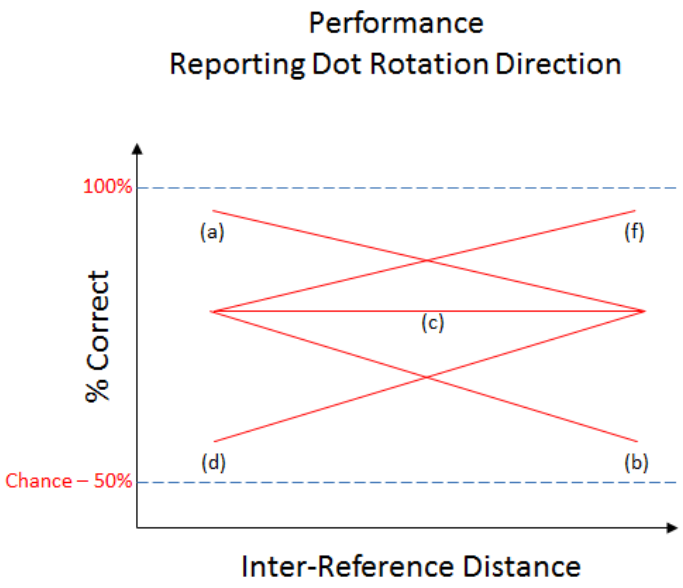


Figure 17: Hypothetical Outcomes of Experiment 3: (a) and (b) inhibitory effect of inter-reference distance on perceptual field strength. (b) and (d) facilitatory effect of inter-reference distance. (c) Independence of field strength from inter-reference distance.

3.5.3 Results

Figure 18 shows performance as a function of the vertical distance between the two neighboring reference frames. Two-way repeated-measures ANOVA indicates that the effect of experimental condition on performance is significant ($F_{3,12} = 53.6$; $p < 0.001$). While the overall effect of distance is found insignificant ($F_{3,12} = 3.2$; $p = 0.088$), there exists a significant interaction between the reference distance and experimental condition ($F_{9,36} = 4.2$; $p = 0.004$). Furthermore, one-way ANOVA shows that the distance between the neighboring squares and the Ternus-Pikler reference disks in fact has a significant effect on performance in the dynamic neighbor experiment condition ($F_{3,12} = 17.3$; $p < 0.001$), but not in the static neighbor condition ($F_{3,12} = 0.1$; $p = 0.878$).

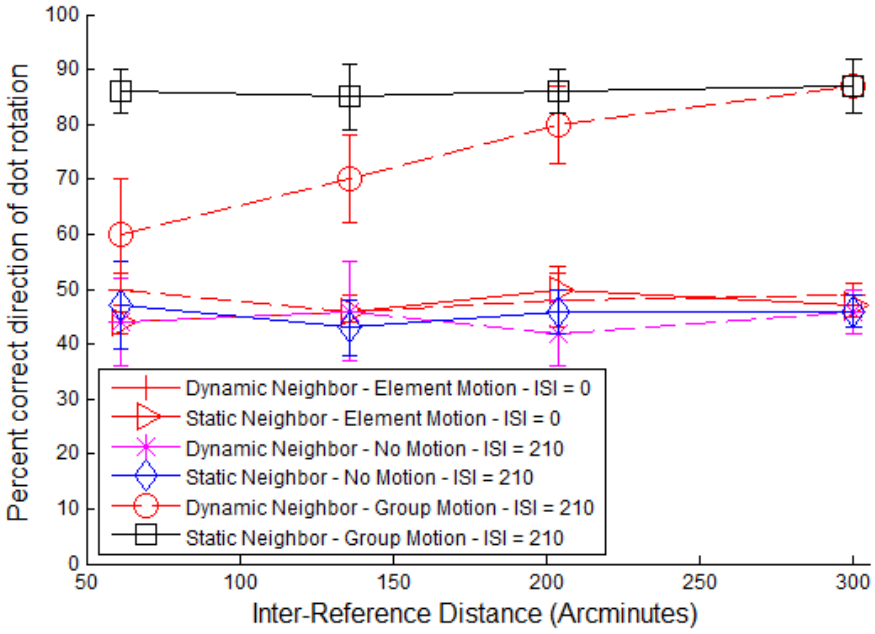


Figure 18: Results from Experiments 3. Percent correct performance in detecting direction of dot rotation, averaged across observers (N=5) and plotted against the center-to-center distance between the disks and squares. Performance for both static and dynamic neighbors is near chance in the absence of group motion. Once group motion is established between Ternus-Pikler disks, subject performance improves. In the case of static neighbor, subject performance remains well above 80%, regardless of the corresponding distance between the neighboring squares and the disks. In the case of dynamic neighbor, however, performance decreases as the distance between the disks and squares is reduced. Error bars correspond to ±1 SEM.

3.5.4 Discussion

The presence of an opposing dynamic reference frame in the neighborhood of the original reference frame interfered significantly with perception of dot rotation. The magnitude of this interference is qualitatively inversely proportional to the distance of the neighboring reference (squares) from the main Ternus-Pikler reference (disks). The magnitude of this interference seems to decrease linearly as inter-reference distance

increases. However, a ceiling effect is observed (beyond 200 arcminutes) as the inter-reference separation approaches approximately four times the diameter of the reference disk.

A static neighboring reference frame, on the other hand, had no significant effect on the perception of dot rotation, even when the dot fell inside the neighboring static objects. These findings indicate that it is in fact the motion of the neighboring squares, and not the squares themselves that serves as a reference frame, capable of interfering with the original field created by the moving disks.

3.6 Experiment 4: Interactions between Same-Direction Neighboring Reference Frames

In Experiment 3, we analyzed the nature of interference between perceptual fields of two dynamic reference frames moving at the same speed, but in the opposite direction. In this experiment, we investigated the characteristics of the net perceptual field produced by two reference frames moving in the same direction with the same speed.

3.6.1 Experimental Methods

The stimulus used in Experiment 4 was similar to that of the dynamic multi-reference case of Experiment 3, with the exception of the direction of motion for the neighboring reference. The three square elements of the neighboring reference were aligned and synchronized with the disks of the original reference to produce two Ternus-Pikler reference frames with identical motion (Figure 19).

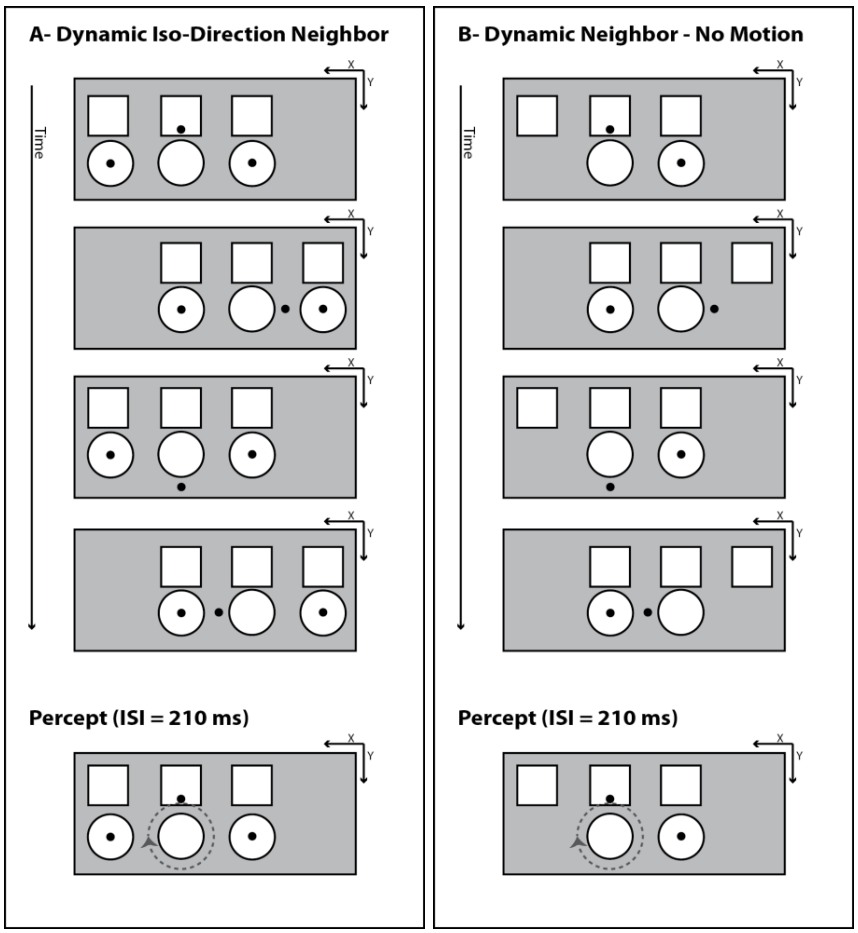


Figure 19: Stimuli used in Experiment 4 and their respective percepts. Stimuli design was similar to that of Experiment 1, with the addition of a neighboring reference frame. Different shapes (disks of 27' radius and squares of 54' sides) were chosen so that elements belonging to the two reference frames remained perceptually different from one another. Element motion and No-motion conditions were included in the experiment, but not shown here. (A) Static Neighbor: Four static squares were introduced above the Ternus-Pikler disks. The distance between the static neighboring set and the Ternus-Pikler reference was varied in the range of 67.86' to 300'. Note that at the minimal separation between the disks and squares, the target dot (4.5' radius) falls inside the boundaries of one of the neighboring squares. (B) Dynamic Neighbor: Three neighboring squares moved in a similar pattern as the Ternus-Pikler disks, but in the opposite direction. All other parameters were identical to the static neighbor condition.

The center-to-center vertical distance between the two neighboring reference frames was varied in the range of 67.86' to 300'. Observers were asked once again to

maintain fixation at the center of the display screen, and to report the perceived direction of dot rotation. Subject population was composed of the same four individuals who took part in Experiment 3. Our informal examination of the stimuli indicated that when the reference frames are perceived to be in element motion, performance remained near chance, as observed in Experiments 1, 2, and 3. Consequently, only the group-motion and no-motion conditions for the Ternus-Pikler disks were examined in this experiment. In four blocks of 200 trials each, subjects were presented with the stimulus in a similar procedure as described in the previous experiments. Both experiment and control conditions were randomized and presented in every block.

3.6.2 Expected Results and Interpretation

Plotting performance as a function of inter-reference distance provides an estimate of the interactions between perceptual fields of neighboring non-retinotopic reference frames (Figure 20). The predicted outcomes “a” and “b” correspond to inhibitory effect, “d” and “f” correspond to facilitatory effect, and “c” corresponds to no effect of inter-reference distance on performance. Linear or non-linear nature of dependence for cases “a”, “b”, “d”, and “f” were also tested. We hypothesize that the absence of motion vectors in the static neighboring reference eliminates the likelihood of any interaction between the two reference frames, and performance remains independent of inter-reference distance (prediction “c”). In the case of dynamic neighboring reference, on the other hand, we hypothesize that the inhibitory interaction between motion vectors of the two aniso-direction reference sets increases as the distance between the elements constituting the two reference frames is decreased. As such, prediction “d” is the more likely outcome.

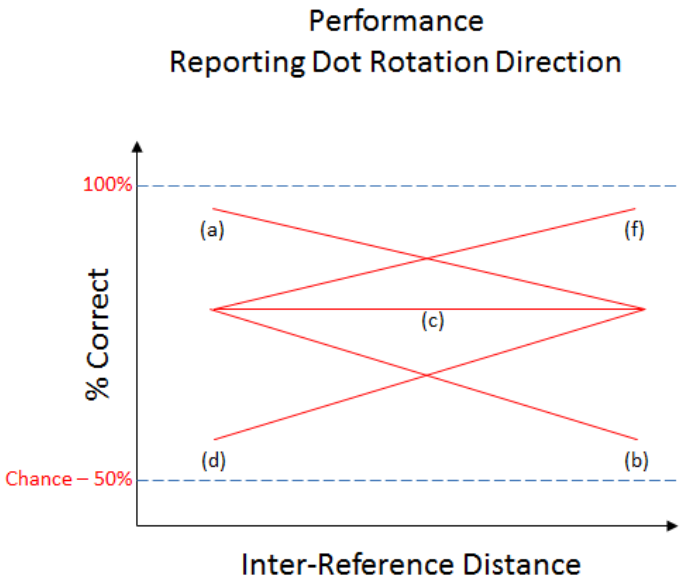


Figure 20: Hypothetical Outcomes of Experiment 4: (a) and (b) inhibitory effect of inter-reference distance on perceptual field strength. (b) and (d) facilitatory effect of inter-reference distance. (c) Independence of field strength from inter-reference distance.

3.6.3 Results

Figure 21 shows performance as a function of the vertical distance between the two neighboring Ternus-Pikler references (disks and squares). Two-way repeated-measures ANOVA shows that both the experimental condition ($F_{1,4} = 12.3$; $p = 0.024$) as well as the distance between the Ternus-Pikler references ($F_{1,4} = 9.5$; $p = 0.027$) significantly affect performance. More specifically, in the no-motion condition, presence of the iso-direction reference frame (squares) in proximal neighborhood of the main reference (disks) significantly improves performance above chance level ($t_{11} = 5.0$; $p < 0.001$). Paired two-sample t-test comparison of performance means, between the same direction neighbor condition of Experiment 4 and the static neighbor condition of

Experiment 3, reveals that the presence of a neighboring reference frame which moves in the same direction as that of the primary reference improves subject performance significantly ($t_{11} = 3.4$; $p = 0.002$).

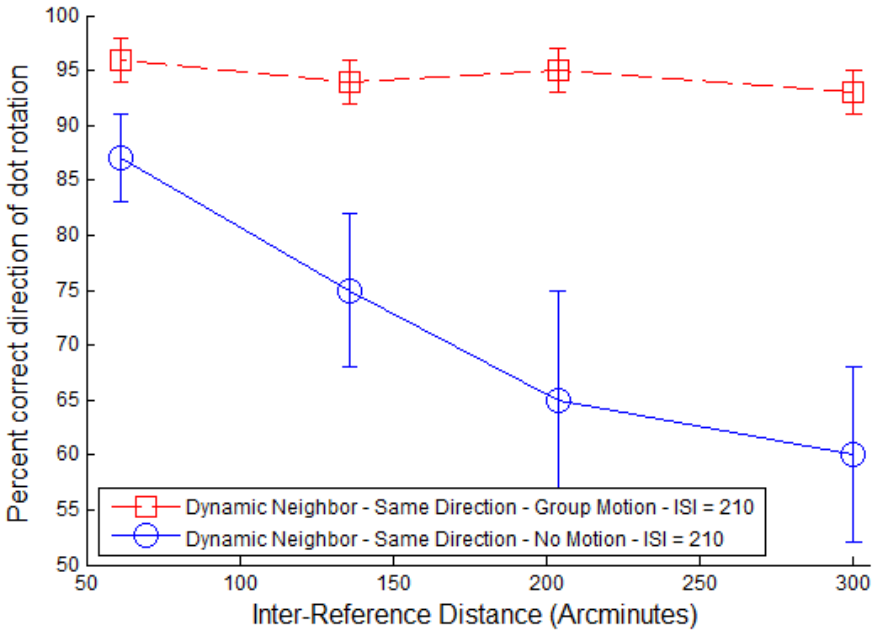


Figure 21: Results from Experiments 4. Percent correct performance in detecting direction of dot rotation, averaged across observers (N=4) and plotted against the center-to-center distance between the two neighboring Ternus-Pikler references (disks and squares). When both sets are perceived to be in group motion, performance remains above 90% regardless of the inter-reference distance. When the Ternus-Pikler disks are in no-motion control condition, however, performance depends on the distance between the two reference frames. Error bars correspond to ±1 SEM.

3.6.4 Discussion

The facilitatory effect of an iso-direction neighboring reference on performance supports the existence of an additive property for non-retinotopic motion fields. These findings are in agreement with those of Experiment 3, where presence of an opposite direction non-retinotopic motion field was found to have an inhibitory effect on performance.

The more surprising finding of this experiment is that, in the presence of the iso-direction neighboring reference frame, a dot rotation percept was induced in the no-motion condition. In our previous experiments, perception of non-retinotopic motion in the target dot was found to be conditional to establishment of group motion between the elements of the main reference (disks). The rotation of the dot was relative to the moving disks; in other words, the dot rotated around the central disk as the central disk was perceived to move. The perceived rotation is analogical to the lunar rotation of the moon around the earth. The results of Experiment 4, however, indicate that presence of a secondary iso-direction reference frame can induce perception of non-retinotopic motion in a given .target, even when the motion signals are removed from its primary reference. Under this condition, the dot does not rotate around the neighboring Ternus-Pikler squares; instead, it rotates around the disks which themselves are not moving and are displaced with respect to the Ternus-Pikler squares. The reference field induced by the squares dynamically shifts the reference from one disk to the other according to the alignment of the disks with respect to the group motion of the Ternus-Pikler squares. The dot is perceived to rotate around the disk which is aligned from frame to frame with the central Ternus-Pikler square. In fact, not only does one perceive a rotation of the dot around a disk, but also the dot perceptually never appears at the center of this reference disk, even though the dots are physically inserted at the center of both disks.

In the absence of primary motion vectors, the influence of the secondary reference frame on a target near the primary reference is qualitatively inversely proportional to the distance between the two reference frames. These findings are in agreement with the

results of Experiment 3, where the magnitude of interference between perceptual fields of opposing reference frames was found to increase as the inter-reference distance decreased. Maintenance of non-retinotopic perception of target motion by the secondary reference frame in absence of primary motion signals can shed new light on processes involved in stabilization of dynamic form under occlusion.

Chapter 4 Temporal Properties of Non-Retinotopic Reference Frames

Visible objects are perceived as unified ensembles of their individual features. The question of binding problem arises since visual features such as form, color, orientation, and motion are said to be coded in physically separate brain modules and pathways (Livingstone & Hubel, 1988; Zeki & Shipp, 1988; Felleman & Van Essen, 1991). How does the visual system correctly combine the features processed in physically segregated cortical areas to achieve a veridical and unified percept of a specific moving object amongst many? In contrast to the classical theories of vision, which viewed the brain as a passive sensory processing unit, more recent theories emphasize the active-constructive nature of the brain and visual perception. There is ample evidence that the temporal structure of both stimulus-driven and top-down processes play a significant role in creation and maintenance of dynamic models of the environment. Indeed the Temporal Binding Hypothesis (TBH) has been proposed as a solution to the binding problem (von der Malsburg, 1981, 1995). The temporal binding hypothesis predicts that, when responding to a common sensory object, functionally specialized neurons from anatomically distinct cortical areas fire their action potentials in temporal synchrony. Temporal synchrony of this assembly of neurons can serve as a solution to the binding problem, as well as a mechanism for figure/ground segregation. The temporal binding hypothesis has received support from neurophysiological (Eckhorn et al., 1988; Engel, Konig, Gray, & Singer, 1990; Livingstone, 1996; Tso & Gilbert, 1988; Tso, Gilbert, & Wiesel, 1986), as well as Electro-EncephaloGram (EEG) studies. The latter suggest that the temporal synchronization of gamma bands (30–90 Hz) activity from independent

areas constitutes a significant component of the binding mechanism (Bertrand & Tallon-Baudry, 2000; Frien, Eckhorn, Bauer, Woelbern, & Kehr, 1994; Gruber, Muller, & Keil, 2002; Muller, Gruber, & Keil, 2000).

Under normal viewing conditions, however, our visual system is highly dynamic. Occlusion and self-occlusion due to object and observer motion can lead to conditions, in which different parts of a given target may appear sequentially or asynchronously. As such, temporal binding model alone cannot overcome the shortcoming of conventional retinotopic theories of vision in explaining how feature binding is achieved under dynamic conditions. There exists ample evidence that visual attributes of a stimulus such as form (Ogmen et al., 2006), luminance (Shimozaki et al., 1999), color (Nishida et al., 2007), size (Kawabe, 2008), and motion (Boi et al., 2009) are computed according to non-retinotopic reference frames. Non-retinotopic reference frames provide a plausible solution to address issues that cannot be explained by purely retinotopic theories of vision. In Experiment 4, we showed that formation of a non-retinotopic reference frame in the neighborhood of a flashing target can lead to perception of coherent motion in the target stimulus. In this study, we examine the temporal characteristics of non-retinotopic reference frames and discuss their potential contribution to address the binding problem.

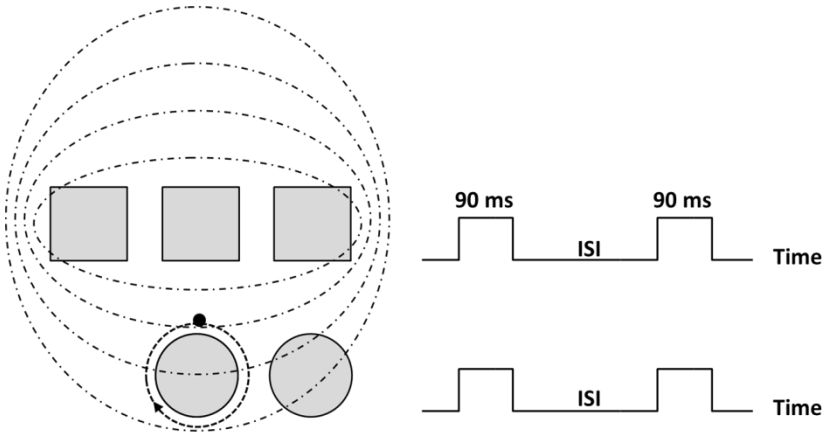
4.1 Experiment 5: Temporal Characterization of the Non-Retinotopic Reference Frames

Experiment 5 was designed to answer two main research questions: i) Does the strength of the non-retinotopic reference frame field effect depend on target-reference synchronization?; and ii) If synchronization plays a role, does the strength of the field effect depend on *absolute* or *relative* timing of target-reference stimulus presentation?

4.1.1 Experimental Methods

The stimulus design in Experiment 5 was similar to the no-motion (two-disk) condition of the dynamic iso-direction multi-reference case in Experiment 4, with one exception: In order to increase the field effect of the neighboring reference frame, two sets of squares were displayed above and below the static disks instead of one (Figure 22). The center-to-center vertical distance between the two neighboring sets of squares and the stationary disks was fixed at 67.86'. All other parameters were chosen to match that of the previous experiment. Observers were asked once again to maintain fixation at the center of the display screen, and to report the perceived direction of dot rotation. Subject population was composed of the same four individuals who took part in Experiments 3 and 4. In blocks of 200 trials each, subjects were presented with the stimulus in a similar procedure as described in the previous experiments. The Ternus-Pikler frame duration was fixed at 90 ms in all conditions. Two different ISI values (270 ms and 450 ms,) as well as several phase shifts were used to explore the temporal characteristics of non-retinotopic reference frames. Different ISI and phase-shift conditions were randomized and presented in each block of the experiment.

A- Perception of Dot Rotation in the Presence of Iso-Direction Neighboring Field



B- Stimulus Design and Target-Reference Phase Shift

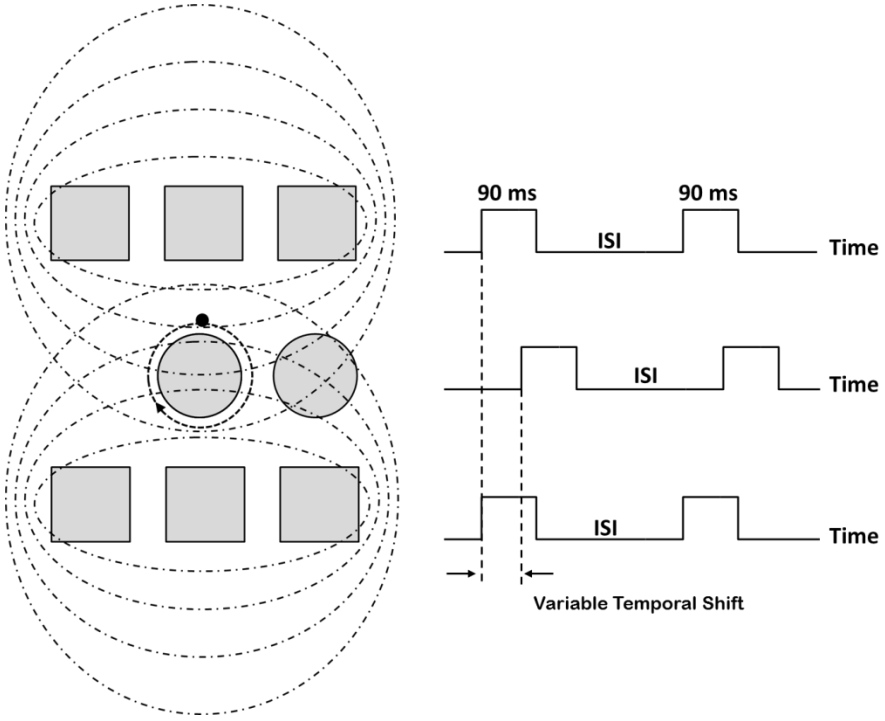


Figure 22: A) The field effect of an iso-direction neighboring reference frame (squares) in the no-motion (two disks) condition in Experiment 4 resulted in perception of dot rotation. Appearance and disappearance of the neighboring squares were synchronized with that of the target Dot and disks. B) A modified version of this stimulus, containing two iso-direction sets of neighboring squares, was utilized in Experiment 5. Two different ISI values (270 ms and 450 ms), as well as several phase shifts, were used to explore the temporal characteristics of

non-retinotopic reference frames.

4.1.2 Expected Results and Interpretation

As depicted in Figure 23, plotting performance against temporal shift can examine whether the non-retinotopic reference frame field strength depends on target-reference synchronization or not. The general shape of the performance function was expected to remain qualitatively the same regardless of the ISI value.

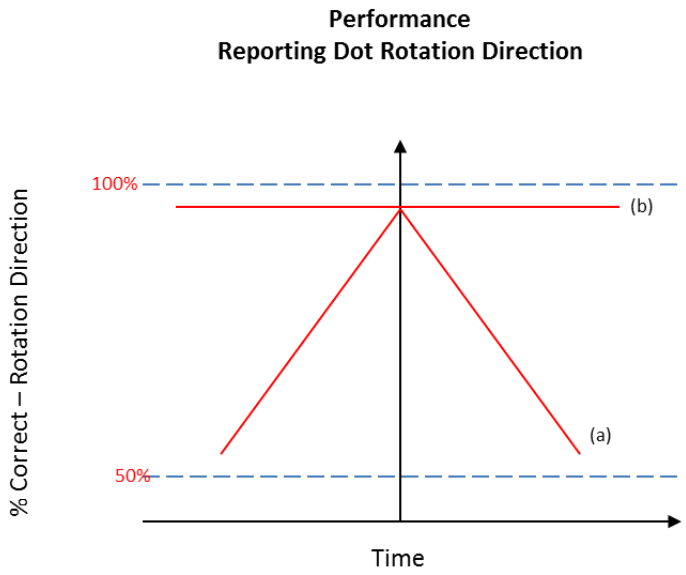


Figure 23: Hypothetical Outcomes of Experiment 5: (a) Inhibitory effect of target-reference asynchrony. (b) Independence of field strength from target-reference synchronization.

Assuming dependence of field strength on target-reference synchronization (Figure 23-a), two different ISI values (270 ms and 450 ms) were chosen to examine whether the field strength depends on *absolute* or *relative* timing of target-reference stimulus presentation. Note that at ISI=270 ms, an absolute time shift results in an equal relative phase shift. However, at ISI = 450 ms the absolute and relative timing is no

longer the same, as the period of the Ternus-Pikler display signal changes from 360 ms to 540 ms.

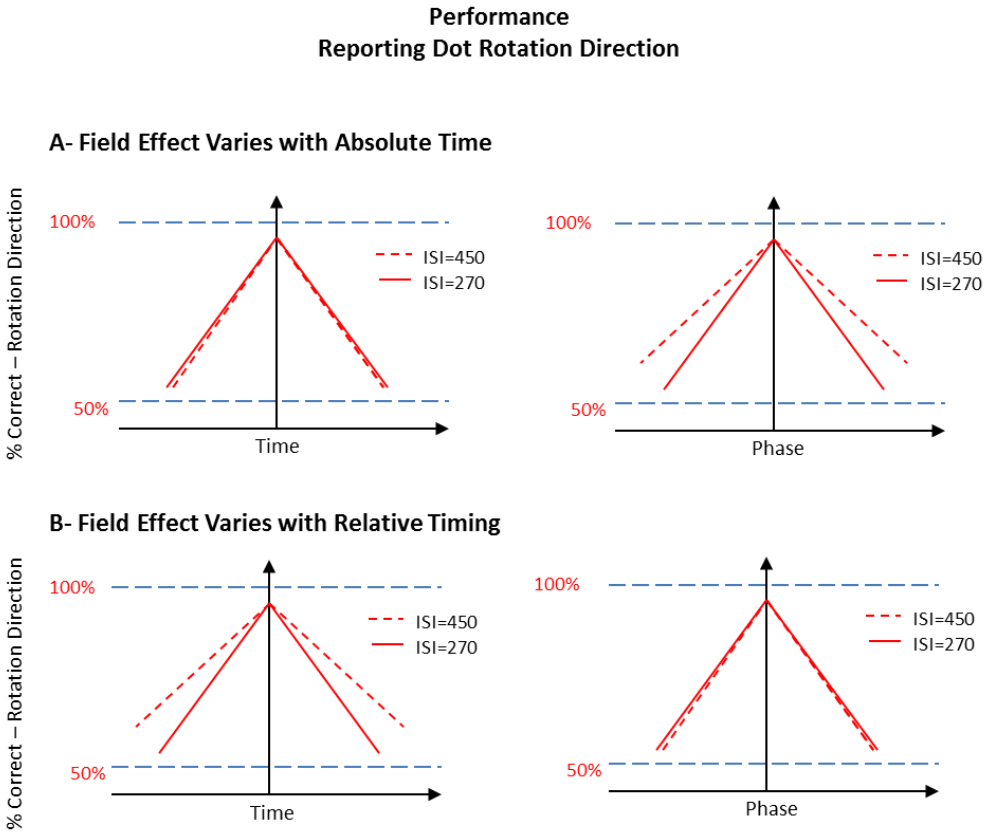


Figure 24: Hypothetical Outcomes of Experiment 5, if the non-retinotopic reference frame field strength depeneds on: (a) absolute time and (b) relative phase shift.

4.1.3 Results

Figure 25 shows the raw results of Experiment 5 averaged across observers. One-way ANOVA indicates that the reference-frame field effect varies with target-reference synchronization regardless of the ISI value ($F_{8,27} = 3.15$; $p = 0.011$ for ISI = 270 ms, and $F_{12,39} = 3.67$; $p < 0.001$ for ISI = 450 ms conditions respectively).

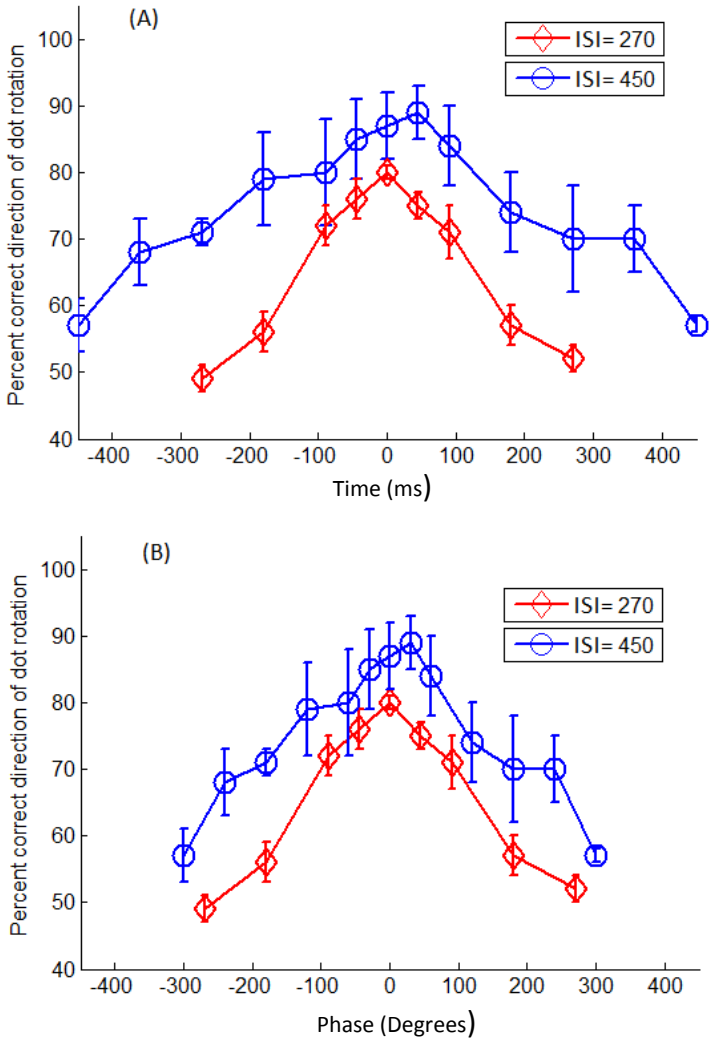


Figure 25: Experiment 5 Raw Results (N=4): Performance plotted against (a) absolute time and (b) relative phase shift.

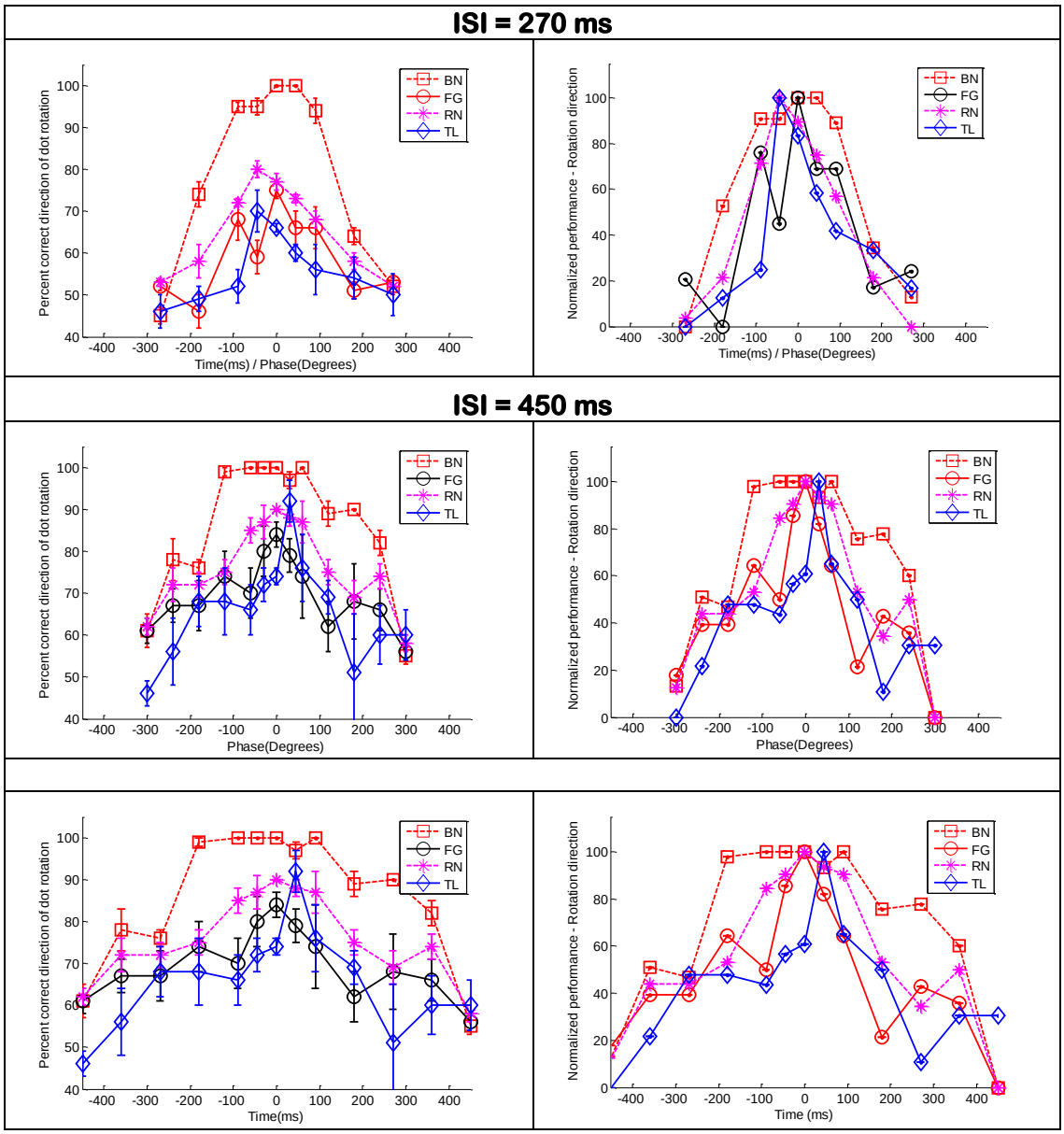


Figure 26: Experiment 5 Individual Performance (left) and Normalized Individual Performance (right) Results: Percent correct dot-rotation direction plotted against absolute time and relative phase. Note that when ISI = 270 ms, time and phase axes overlap.

Individual performance results, however, showed that one subject performed significantly better than others (Figure 26 – left column). In order to prevent domination

of average results by one subject, individual performance results were normalized across all subjects. The normalization is justified since our main interest is the dependence of performance on time vs. phase shift, rather than absolute magnitude of performance. Individual subject performance results were normalized according to the following equation:

$$Normalized_x = 100\frac{[x - \min(x)]}{[\max(x) - \min(x)]}, \quad (4.1)$$

$$where\ x\ is\ a\ vector\ containing\ an\ individual\ subjects'\ performance\ values$$

Normalized results were fitted with Gaussian curves using Matlab *normfit* function (Table 4.1), and the fitted results were plotted using the *normpdf* function (Figure 27). Paired two-sample t-test comparison of variance between the fitted results of individual subjects indicates that ISI effect is significant in both absolute time (t_3 = -20.06; $p < 0.001$; d = 1.975), as well as phase domain (t_3 = -3.32; p = 0.001; d = 0.791). The Cohen's d, however, is much smaller in phase domain, signifying the importance of relative timing or phase shift on the strength of the non-retinotopic field effect.

Table 1: Mean and Standard Deviation Values for Individual Gaussians

	ISI = 270		ISI = 450			
	Phase/Time		Phase		Time	
Subject	mean	sd	mean	sd	mean	sd
BN	0.63	98.54	1.09	127.97	1.63	191.95
FG	11.03	110.78	-15.56	125.59	-23.35	188.38
RN	-7.31	82.22	-4.15	122.89	-6.22	184.34
TL	20.79	95.11	13.42	129.17	20.13	193.75

* All shown time values are in milliseconds and phase values are in degrees

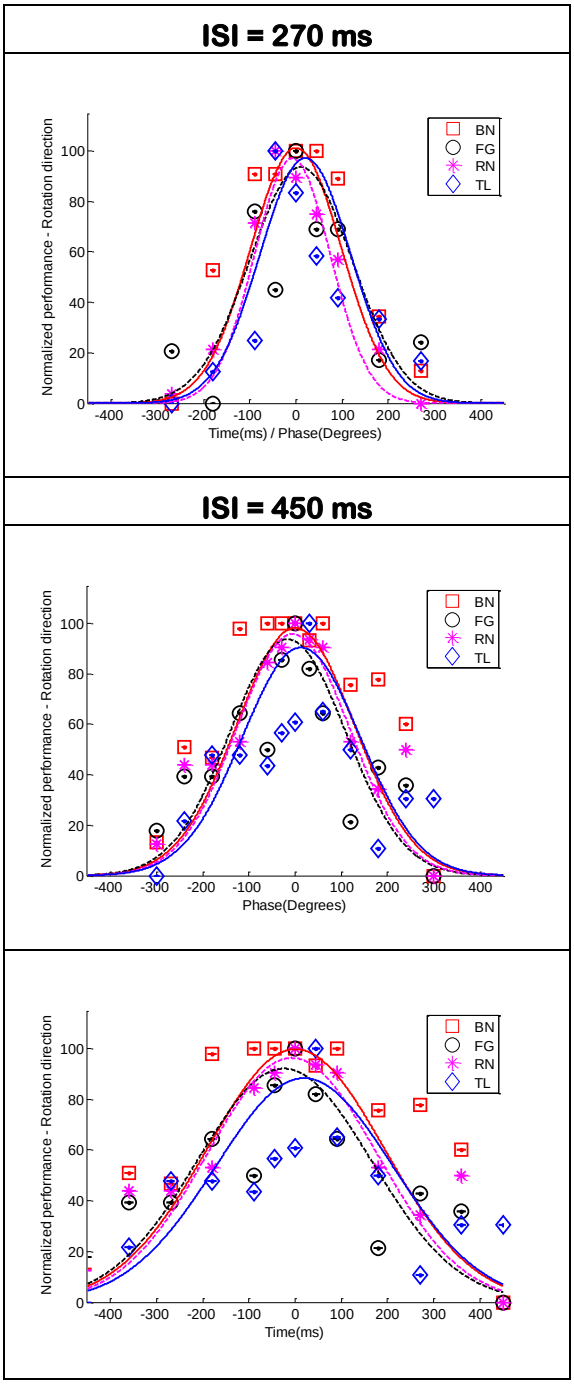


Figure 27: Experiment 5 Results Fitted with Gaussian Curves: Performance plotted against absolute time and relative phase shift reveals that relative phase shift

between target and reference stimuli presentation is a more significant factor determining the strength of the non-retinotopic reference frame field effect.

4.1.4 Discussion

These results indicate that the field effect of a non-retinotopic reference frame on a given target is qualitatively proportional to the target-reference temporal synchrony. Furthermore, the strength of the non-retinotopic field effect depends on both absolute and relative timing of target-reference stimulus presentation, with relative timing or phase shift having a stronger effect. These results collectively suggest that common local motion vectors can serve as a non-retinotopic reference frame to bind targets, which appear in brief temporal asynchrony. The inherent flexibility of this type of temporal binding can potentially resolve the aforementioned issues arising from purely retinotopic processing of occlusion and self-occlusion. In other words, the common motion of an object can serve as a reference frame to bind the parts of the object that come in view sequentially or asynchronously. However, as the magnitude of the target-reference relative asynchrony increases, the effect of the reference frame field reduces.

Chapter 5 Effects of Endogenous Attention on Reference Frame Field Strength

In the previous experiments, observers were asked to focus their attention on Ternus-Pikler elements. In our informal observations of the stimuli presented for long durations, we noticed that the allocation of attention to other parts of the display could alter the percepts. In addition, allocation of attention has been found to influence the likelihood of perceiving group motion in Ternus-Pikler displays (Aydin, Herzog, & Ogmen, 2011). Since formation and maintenance of non-retinotopic reference frames depend critically on the perception of group motion among the elements of the Ternus-Pikler reference, we hypothesized that diversion of attention in our experiments should influence the strength of non-retinotopic reference fields. In Experiment 5, we studied the effect of attention on non-retinotopic reference frames.

5.1 Experiment 6: Attention Modulates Non-Retinotopic Reference Frame Field Effect

5.1.1 Experimental Methods

The stimulus used in Experiment 5 was identical to that of the static multi-reference case of Experiment 3. In order to study the effects of attention on reference frame strength, participants were instructed to focus their attention on the two central elements in the presented set of four static squares. The task was once again to report the perceived direction of dot rotation. Fixation was maintained at the center of the display screen. Once again, the center-to-center vertical distance between the neighboring

squares and the Ternus-Pikler reference disks was varied in the range of 67.86' to 300', and subject performance was measured. Subject population was composed of the same four individuals that took part in Experiments 3 and 4. With subjects attending the neighboring static reference, responses were collected in four blocks of 150 trials each, in accordance with procedures discussed in the previous experiments.

5.1.2 Expected Results and Interpretation

We hypothesize that allocation of attention to the neighboring static reference frame instead of the main reference significantly attenuates performance. As such, the expected performance should be lower than that found in the static condition in Experiment 3.

5.1.3 Results

Figure 28 shows performance as a function of the vertical distance between the neighboring square objects and the Ternus-Pikler reference disks. The data displayed here includes the static neighbor results obtained in Experiment 3 as well. Collectively, four experimental conditions were included based on perceived motion of the Ternus-Pikler disks (group or element) and the locus of attention (Ternus-Pikler disks or neighboring squares). Two-way repeated-measures ANOVA shows that the experimental condition ($F_{3,9} = 30.8$; $p = 0.002$), but not the inter-reference distance ($F_{3,9} = 0.12$; $p = 0.947$) is significant. Performance remains near chance in the element motion condition, regardless of locus of attention ($t_{11} = -0.43$; $p = 0.334$ and $t_{11} = 1.57$; $p = 0.068$ respectively for attending disks or squares). When the element motion conditions are removed from the obtained results, two-way repeated-measures ANOVA shows that attending the static

neighboring reference instead of the Ternus-Pikler disks significantly attenuates the average performance ($F_{1,3} = 13.7$; $p = 0.034$).

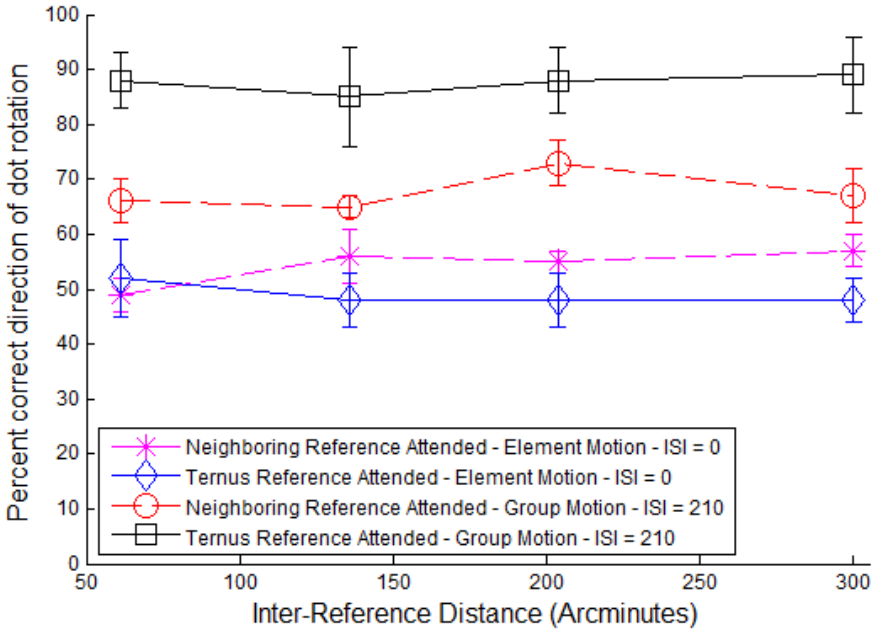


Figure 28: Results from Experiments 6. Percent correct performance in detecting direction of dot rotation, averaged across observers (N=4) and plotted against the center-to-center distance between the Ternus-Pikler reference disks and the neighboring set of static squares. Performance is near chance in the absence of group motion. Once group motion is established between Ternus-Pikler disks, performance improves. When the Ternus-Pikler disks are in group motion and attended, performance remains well above 80%. Average performance, however, attenuates significantly when the neighboring static squares are attended instead of the Ternus-Pikler disks. Inter-reference distance has no significant effect on performance, regardless of the perceived reference motion or the locus of attention. Error bars correspond to ±1 SEM.

5.1.4 Discussion

The results of Experiment 6 collectively indicate that diversion of attention from the main reference significantly attenuates the strength of motion-based non-retinotopic reference frames. These findings emphasize the role of top-down perceptual processes in

establishment and maintenance of non-retinotopic reference frames, and support earlier reports (Aydin et al., 2011) on the significant role of attention in modulation of spatio-temporal grouping. Aydin et al. showed that diverting attention away from Ternus-Pikler elements reduces the probability of group motion percept. Since establishment and maintenance of non-retinotopic reference frames depend on perception of group motion, diverting attention away from the Ternus-Pikler disks (to the static neighboring squares) reduces the strength of the perceptual field induced by the non-retinotopic Ternus-Pikler reference frame. Moreover, it was sufficient to divert attention at the closest distance to reduce the effect, and spreading attention further away in space did not cause any additional drop in performance.

Chapter 6 Visual Masking Experiments on Information Transfer from Retinotopic to Non-Retinotopic Representation

Due to the movements of the observer and those of objects in the environment, retinotopic representations are highly unstable during ecological viewing conditions. The phenomenal stability of our perception suggests that retinotopic representations are transformed into non-retinotopic representations. It remains to show, however, which visual processes operate under retinotopic representations and which ones operate under non-retinotopic representations. Visual masking refers to the reduced visibility of one stimulus, called the target, due to the presence of a second stimulus, called the mask (Bachmann, 1994; Breitmeyer & Ogmen, 2006). It has been used extensively to study dynamic aspects of visual perception. Previous studies using Saccadic Stimulus Presentation Paradigm (SSPP) suggested both retinotopic and non-retinotopic bases for visual masking. In order to understand how the visual system deals with retinotopic changes induced by moving targets, here we investigated i) the retinotopy of visual masking and ii) the fate of masked targets under conditions that do not involve eye movements. We have developed a series of experiments based on a radial Ternus-Pikler display. In this paradigm, the perceived Ternus-Pikler motion is used as a non-retinotopic reference frame to pit retinotopic against non-retinotopic visual masking hypothesis. Our results indicate that both metacontrast and structure masking are retinotopic. We also show that, under conditions that allow observers to read-out effectively non-retinotopic feature attribution, the target becomes visible at a destination different from its

retinotopic/spatiotopic location. We discuss the implications of our findings within the context of ecological vision and dynamic form perception.

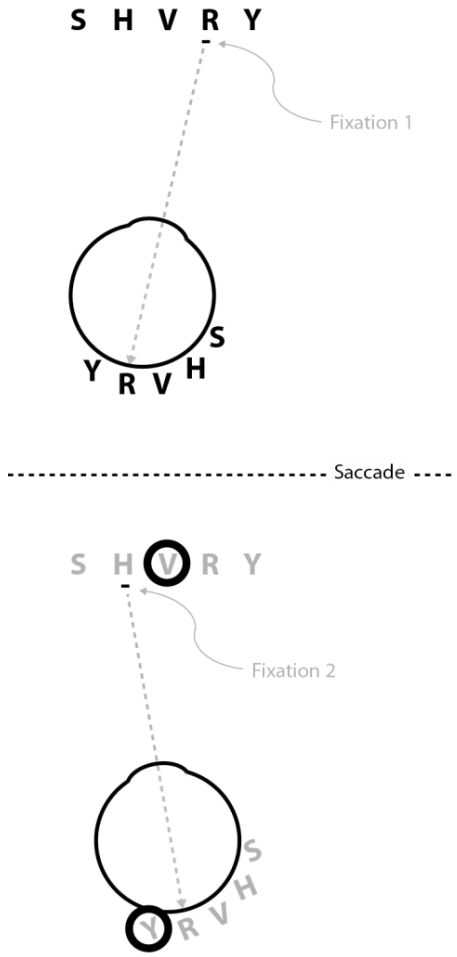


Figure 29: Saccadic Stimulus Presentation Paradigm (SSPP) used by Davidson et al. (1973). The observer makes a saccade from the first to the second fixation. Target stimuli, consisting of five letters are presented briefly just before the initiation of the saccade. A mask stimulus is presented after the saccade. The light gray letters at the bottom are shown to highlight the relative position of the mask with respect to the targets. In the actual stimulus, letters were only presented before the saccade. As one can see from the figure, the mask surrounds letter V according to spatiotopic coordinates and letter Y according to retinotopic coordinates. The non-overlapping ring mask shown here corresponds to the metacontrast condition. The experiments also had an overlapping pattern to examine pattern masking by structure.

Saccadic eye movements constitute a major source for retinotopic instability. However, during these eye movements, the world appears phenomenally stable suggesting that retinotopic shifts caused by the saccades are either dismissed or compensated by the visual system. Theories suggesting the dismissal solution maintain that very little information is kept from one saccade to another and vision starts *tabula rasa* after each saccade. Theories suggesting complete compensation solution maintain that all information is remapped across the saccade by taking into account the global shift caused by the saccade. Theories that take intermediate positions between these two extremes have also been proposed (rev., Bridgeman, van der Heijden, & Velichkovsky,, 1994). In general, compensation theories rely on three mechanisms: (i) prior to the initiation of the saccade, retinotopic information is stored in memory, (ii) during and after the saccade, retinotopic information is suppressed to prevent inappropriate integration of pre- and post-saccadic images, (iii) after the saccade, the new image is integrated with the contents of the memory by taking into account the retinotopic shift caused by the saccade. Because saccadic shifts take in general few tens of milliseconds, sensory (iconic) memory[1] and backward masking[2] have been viewed as the major candidates to carry out the memorization and suppression tasks, respectively.

Several studies investigated whether sensory memory, masking, and information integration occur in retinotopic or non-retinotopic coordinates across saccadic eye

[1] Sensory (iconic) memory is a visual storage mechanism with relatively high capacity and a relatively short time span.

[2] Backward masking refers to the reduction in the visibility of a target stimulus caused by a mask stimulus that follows the target in time. When the mask surrounds but does not spatially overlap with the target, it is called metacontrast. When the mask spatially overlaps and shares structural similarities to the target, it is called backward structure masking (Bachmann, 1994; Breitmeyer & Ogmen, 2006).

movements[3]. Saccadic Stimulus Presentation Paradigm (SSPP) has been the classical experimental paradigm for these studies. (Davidson et al., 1973; Irwin, 1991; Knapen et al., 2009; McRae et al., 1987; Melcher & Colby, 2008; Melcher & Morrone, 2003). Figure 29 shows the SSPP paradigm used by Davidson et al. (1973) to investigate retinotopic versus non-retinotopic bases of backward masking.

The observer is asked to make a saccade from one fixation point to a second one. Target stimuli (five letters) are presented briefly before the saccade, followed by a mask stimulus (either a non-overlapping ring, as in Fig. 1, or an overlapping pattern) presented after the saccade. As depicted in Figure 29, the mask stimulus surrounds (or covers) different letters according to retinotopic and non-retinotopic (spatiotopic) coordinates. By measuring which of the two letters is suppressed from perception, one can infer whether this mask operates in retinotopic or non-retinotopic coordinates. Davidson et al. (1973) reported that the mask suppressed the letter that shared its retinotopic coordinates, but appeared to occupy the same position as the letter that shared its spatiotopic coordinates. They suggested the existence of retinotopic visible persistence at which trans-saccadic masking occurs, and a spatiotopic sensory memory at which trans-saccadic integration occurs. In a follow-up study, McRae et al. (1987) reported not only retinotopic but also spatiotopic masking. They suggested that the transition from retinotopic to spatiotopic representations takes time and the reason Davidson et al. (1973) did not find evidence for spatiotopic masking could be the relatively shorter Inter-Stimulus Interval (ISI) used by Davidson et al. (ca. 80 ms) compared to the ISI used in their study (153 ms). That

[3] White used smooth pursuit eye movements to study retinotopic versus non-retinotopic aspects of visual masking (White, 1976). He reported spatiotopic masking. However, a subsequent study where eye movements were monitored showed that masking during pursuit was retinotopic and not spatiotopic (Sun & Irwin, 1987).

masking is retinotopic at short ISIs was also confirmed by Irwin et al. (1988). These authors also presented evidence for spatiotopic memory integrating information across saccades. However, their data suggested that spatiotopic integration of information was rather abstract depending on position and identity information rather than detailed image-like fusion of trans-saccadic stimuli (see also van der Heijden, Bridgeman, & Mewhort 1986).

The observation of shifts of neuronal receptive fields in the direction of intended saccades (Duhamel, Colby, & Goldberg, 1992) generated a renewed interest for the problem of visual stability across saccades from this perspective (rev. Melcher & Colby, 2008; Wurtz, 2008). These "remapping of receptive fields" have been associated with shifts in the perceived positions of peri-saccadically presented targets (rev., Ross, Morrone, Goldberg, & Burr, 2001). Pisapia, Kaunitz, & Melcher (2010) suggested that these shifts, in turn, can help a target stimulus escape from making. Moreover, they have also presented evidence for spatiotopic masking for ISIs shorter (48 ms) than the ISIs reported in previous studies. Hunt and Cavanagh (2011) presented a brief target before the saccade followed by a long-duration mask that turned on before the saccade and remained on after the saccade until the subject responded. With this paradigm, they showed masking when the mask was presented at the *post-saccadic retinotopic coordinates* of the location where the target was presented. Taken together, these studies paint a complex picture for the retinotopy of masking. Part of the reason for this complexity may be due to the fact that many of the studies used different types of target, mask pairs and widely different parameters. It is known that masking is not a unitary phenomenon (Breitmeyer & Ogmen, 2006) and the differences between the studies may

be due to differences in the types of masking functions and mechanisms evoked in different studies. Notwithstanding this issue, these studies show that SSPP provides a powerful method for exploring retinotopy of visual masking across saccades. However, this paradigm involves eye-movement related processes, such as saccadic suppression and efference copy, and cannot be employed to study retinotopy of visual masking independent of eye movements.

6.1 Retinotopy of Visual Masking in the Absence of Eye Movements

More recently, an alternative method for exploring non-retinotopic processing based on the Ternus-Pikler paradigm has been proposed (Boi et al., 2009; Ogmen et al., 2006). The Ternus-Pikler display is an apparent motion stimulus, introduced by Gestalt psychologists about a century ago, and employed extensively since then to study the spatio-temporal aspects of human vision (Petersik & Rice, 2006; Pikler, 1917; Ternus, 1926). Figure 30 shows how the Ternus-Pikler stimulus has been adopted for studying non-retinotopic processing of stimulus features (Ogmen et al., 2006). The basic Ternus-Pikler paradigm (Figure 30-A) consists of two display frames separated by a blank frame presented for the duration of ISI. The two display frames are identical, except that all elements in Frame 2 are shifted by one inter element distance with respect to the elements in the first frame.

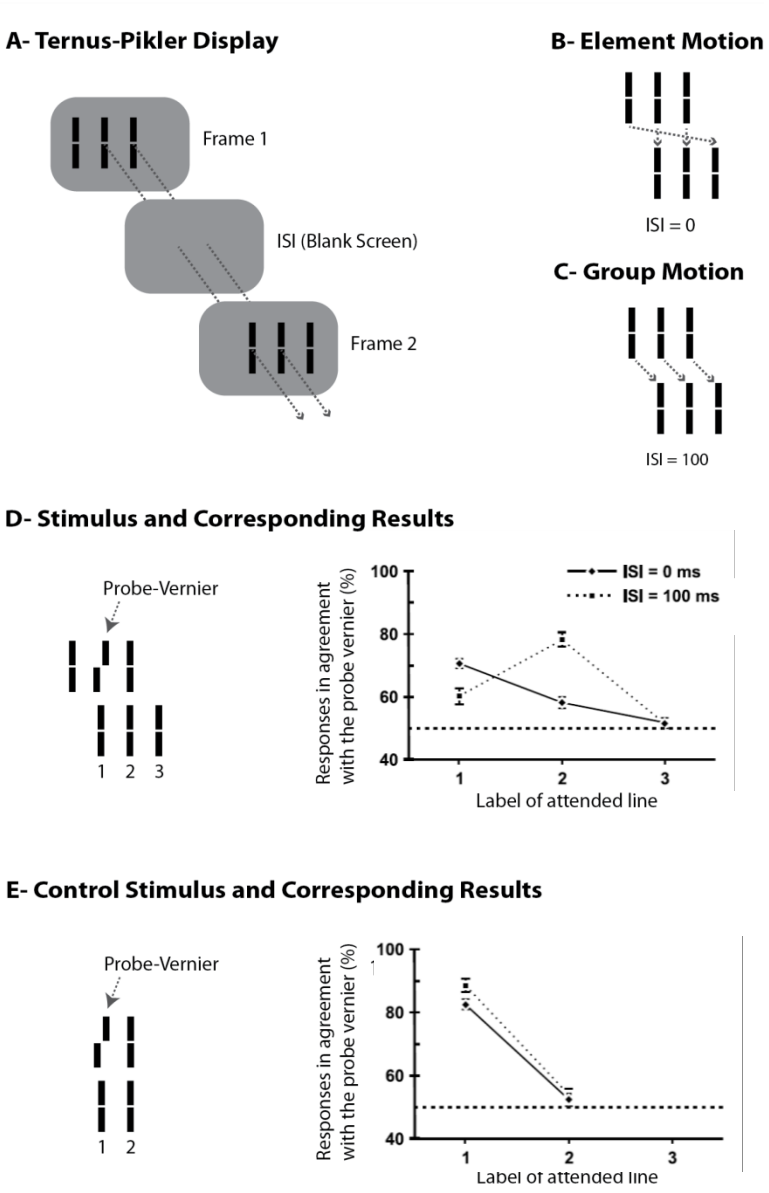


Figure 30: Ternus-Pikler paradigm for exploring non-retinotopic feature processing. (A) Ternus-Pikler Display: two display frames are separated by a blank interval called Inter Stimulus Interval (ISI). The two display frames are identical, except that all elements in Frame 2 are shifted by one inter element distance with respect to the elements in the first frame. (B) Element Motion: For short ISIs (e.g., 0 ms) observers perceive the leftmost element in Frame 1 to be moving to the position of the rightmost element in Frame 2. In this case, no motion is perceived for the other two elements. (C) Group Motion: For long ISIs (e.g., 100 ms) observers perceive all elements to be moving as a group. (D) Stimulus and the Corresponding Results: A Ternus–Pikler display presented with an ISI of either 0 or 100 ms. The central element in the first frame included a small offset called the "probe-vernier". Observers attended to one of the elements of the second frame

labeled as 1, 2, or 3. (E) Control Stimulus and the Corresponding Results: Only the elements that overlapped in the two Ternus-Pikler frames were shown, i.e., the leftmost element of the first and the rightmost element of the second frame of the stimulus shown in (D) were not displayed. No motion percept was elicited (Adapted from Öğmen et al., 2006).

Depending upon the ISI duration, two different types of motion are perceived (Pantle & Picciano, 1976). For short ISIs (e.g., 0 ms) observers perceive Element Motion, in which the leftmost element in Frame 1 is perceived to be moving to the position of the rightmost element in Frame 2 (Figure 30-B). In this case, no motion is perceived for the other two elements. For long ISIs (e.g., 100 ms) observers perceive Group Motion, in which all elements are perceived to be moving together as a group (Figure 30-C). In order to apply this paradigm to study feature processing, a simple spatial feature called a vernier offset can be inserted into the central element of the first frame (Figure 30-D). When fixation is maintained in a typical Ternus-Pikler episode, leftmost element of Frame 2 is always displayed at the same retinotopic location as that of the central element of Frame 1. Purely retinotopic hypotheses consequently predict that features of the central element in Frame 1 should be integrated into the leftmost element of Frame 2 for all ISI values within the window of temporal integration. For instance, such retinotopic hypotheses predict that performance should be well above chance when subjects are asked to report the direction of the probe vernier perceived in the leftmost element in Frame 2, and near chance for the other elements in Frame 2. However, it was shown that subject performance in reporting direction of vernier offset depends on the ISI value. More specifically, when group motion is established between the Ternus-Pikler elements (ISI = 100), performance is well above chance when subjects report the vernier offset perceived in the central element in Frame 2, and near chance for other elements (Figure

30-D). On the other hand, in the case of element motion (ISI = 0), performance is higher when subjects report the vernier offset perceived in the leftmost element in Frame 2. The illusory attribution of the vernier offset also depends critically on the elicitation of a motion percept. If the leftmost line of the first frame and the rightmost line of the second frame are omitted (Figure 30-E), no apparent motion is induced since the remaining elements spatially overlap. In this control display, percentage of responses in agreement with the probe-vernier is high only for the element labeled 1 and at chance level for element 2 for both ISIs. These results collectively indicate that feature processing and attribution between elements of consecutive Ternus-Pikler display frames is governed according to motion-induced grouping; i.e. according to the dashed arrows in Figure 30-B and 2-C. In other words, perceived motion of the Ternus-Pikler elements serves as a non-retinotopic reference frame for feature processing and attribution between elements of the Ternus-Pikler display frames. In the present study, we use a similar Ternus-Pikler paradigm to *i)* probe retinotopic and non-retinotopic bases of visual masking, and *ii)* assess non-retinotopic perception during masking. We utilized the above mentioned experimental technique to study visual masking in the absence of saccadic eye movement. The experiments discussed in this section were collectively designed to achieve the following aim:

> ***Specific Aim 4:*** To determine the factors controlling the transfer of information from retinotopic to non-retinotopic representation in the presence of deblurring mechanisms such as visual masking.

6.2 Experiment 7: Retinotopy of Metacontrast Masking

In this experiment, we utilized a radial Ternus-Pikler display to study retinotopy of metacontrast masking in the absence of eye movement. Experimental procedures and results are discussed in the following sections.

6.2.1 Experimental Methods

Two display frames, each of which contained two discs and a central square aligned on the perimeter of an invisible circle centered at the fixation, were displayed sequentially to create perception of radial Ternus-Pikler apparent motion (Figure 31-A and 31-B). The radius of this virtual circle was fixed at 2.5 degrees in all experiments. The target-mask combination shown in Figure 31-C was displayed at variable Stimulus Onset Asynchronies (SOAs). The target was predictably presented at the center of the square in the first frame of the Ternus-Pikler sequence, and subjects were asked to attend and report the location of the missing corner on the black target diamond (left/right). Depending upon the spatial location of the mask within Frame 2, retinotopic and non-retinotopic masking effects were distinguished. Figure 31-A displays the case of *Retinotopic Mask Condition*, where the mask in Frame 2 was presented at the same retinotopic location as that of the target diamond in Frame 1. Note that in the absence of eye movements, retinotopic and spatiotopic masking conditions are equivalent. Figure 31-B, on the other hand, depicts the case of *Non-Retinotopic Mask Condition*. In this case, the mask was displayed in the central square of Frame 2. The two squares presented in Frames 1 and 2 of the Ternus-Pikler sequence correspond to one another only when group motion is established.

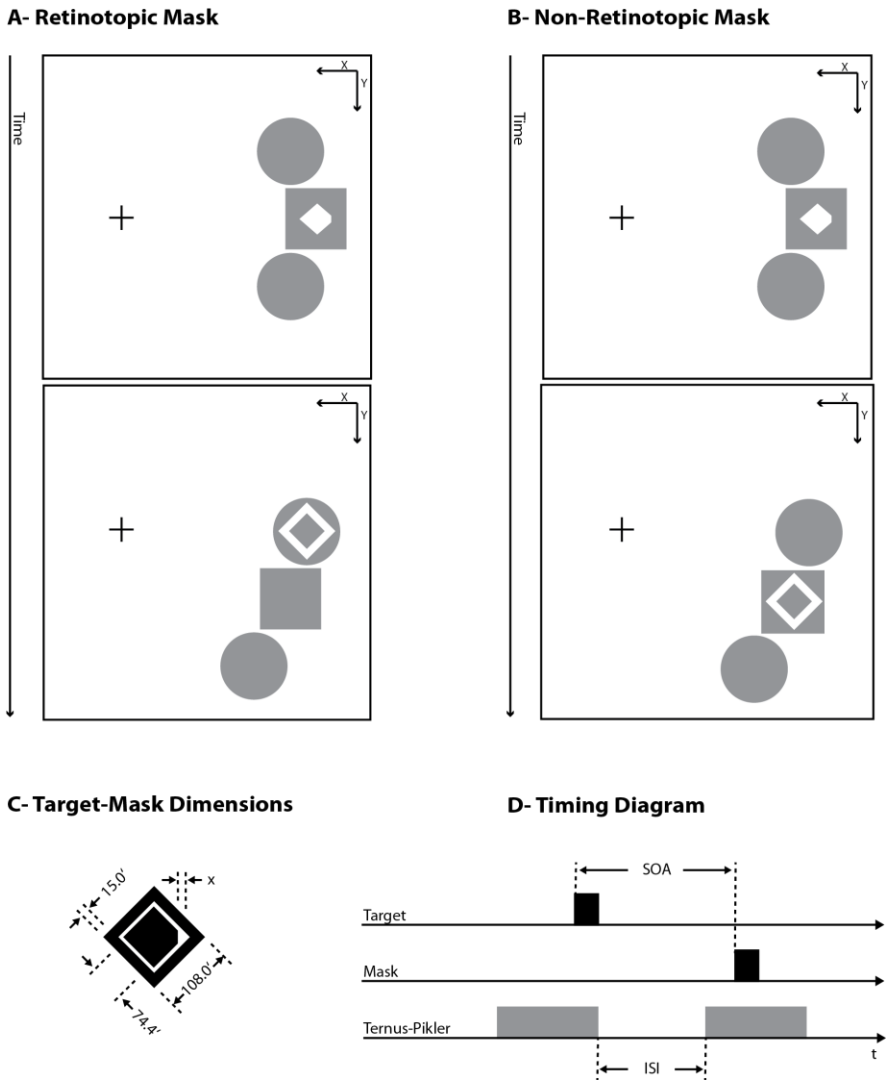


Figure 31: Stimuli and Respective Parameters for Experiment 1: Two display frames, each of which contained two disks and a square, were displayed sequentially to create perception of radial Ternus-Pikler apparent motion. The blank ISI frame is not displayed in this figure for the sake of simplicity. Subjects were asked to report the location of the missing corner on the black target diamond, shown at the center of the middle square in the first frame. (A) Retinotopic Mask Condition: The mask was displayed in Frame 2, at the same spatial location as that of the target diamond in Frame1. (B) Non-Retinotopic Mask Condition: The mask was displayed in the central square of Frame 2, which corresponded with the central square of Frame1, only when disks were perceived to be in group motion. (C) Spatial Parameters of the Target and Mask: Variable "x" represents the size of the probe gap, which was varied to meet individual subject threshold requirements. (D) Timing Diagram: ISI value was fixed (0 or 40 ms) per block, and the target predictably appeared just before the ISI. Mask presentation time, however, was randomized from trial to trial in order to allow for different ISI-SOA combinations per block.

In addition to the retinotopic and non-retinotopic mask experiment conditions, two control conditions were included in this experiment. In the *Static Control Condition*, masking functions were obtained for individual subject in the absence of Ternus-Pikler motion. Under this condition, the Ternus-Pikler elements remained visible throughout the experiment at the same spatial location as that of Frame1 in Figure 31-A or 31-B. In the *No-Mask Control Condition,* target was shown in the absence of the mask.

Spatial parameters of the target and mask are displayed in Figure 31-C. Variable "x" represents the size of the probe gap, which was varied in the range of 12' to 25' to meet each individual subject's masking threshold requirements. Figure 31-D displays the timing diagram of a typical trial. The ISI was fixed (0 or 40ms) for each experimental block, and the target always appeared just before the ISI. Target and mask stimuli were presented for 10 ms, each. Mask onset time was randomized from trial to trial to allow for different ISI-SOA combinations per block. As it can be seen in Figure 3D, the Ternus-Pikler ISI limits the shortest masking SOA that can be used. Therefore, eccentricity, background luminance, Ternus-Pikler element shapes, and target/mask/disk contrasts were chosen in such a way that Ternus-Pikler group motion was perceived by all observers at a relatively short ISI (40 ms), while strong masking effect was observed at the corresponding SOA (50 ms). Ternus-Pikler radial motion (upward or downward) was also randomized from trial to trial. In a two-alternative forced-choice design, three naïve observers as well as one of the authors reported the perceived location of the missing corner of the target diamond (left/right).

6.2.2 Expected Results and Interpretation

Note that in the non-retinotopic mask condition (Figure 31-B), the target and mask always stimulate distinct retinal areas. Retinotopic theories predict no masking effect for this condition, regardless of stimulus timing and Ternus-Pikler grouping. However, non-retinotopic theories predict that in such a case, masking effect follows stimulus timing and the perceived Ternus-Pikler motion. As such, if the Ternus-Pikler elements are perceived to be in element motion, the non-retinotopic prediction is same as the retinotopic hypothesis. However, when Ternus-Pikler elements are perceived to be in group motion, the non-retinotopic hypothesis predicts masking effect for non-retinotopic mask condition instead. Figure 32 summarizes the respective predictions of retinotopic and non-retinotopic hypotheses, based on the perceived motion grouping of the Ternus-Pikler disks.

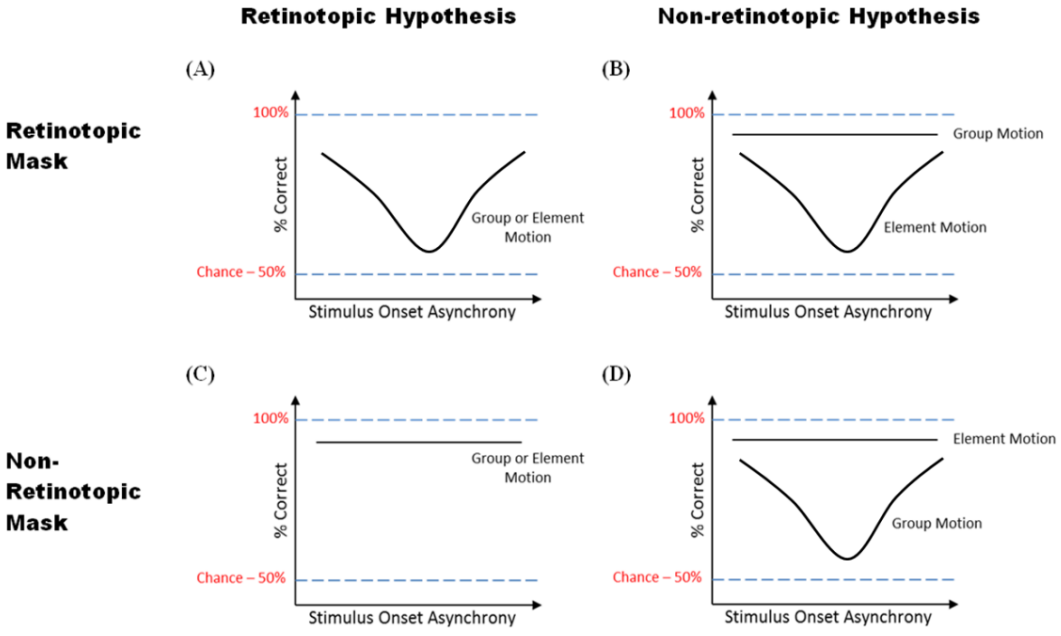


Figure 32: Predictions of Retinotopic and Non-Retinotopic Theories for Experiment 1: Panels (A) and (C) depict predictions of Retinotopic theories. Masking effect is expected only in retinotopic mask experiment condition, regardless of the

perceived Ternus-Pikler motion (group or element). Panels (B) and (D) depict predictions of Non-Retinotopic theories. Masking effect for each experiment condition is expected to depend on perceptual grouping of Ternus-Pikler disks.

6.2.3 Results

Figure 33 shows performance as a function of the SOA for the static control condition. Comparison of masked and no-mask conditions shows significant masking effect (t_5 = -5.14; p = 0.002). As expected, metacontrast masking function dips at SOA = 40 ms, indicating type B masking function.

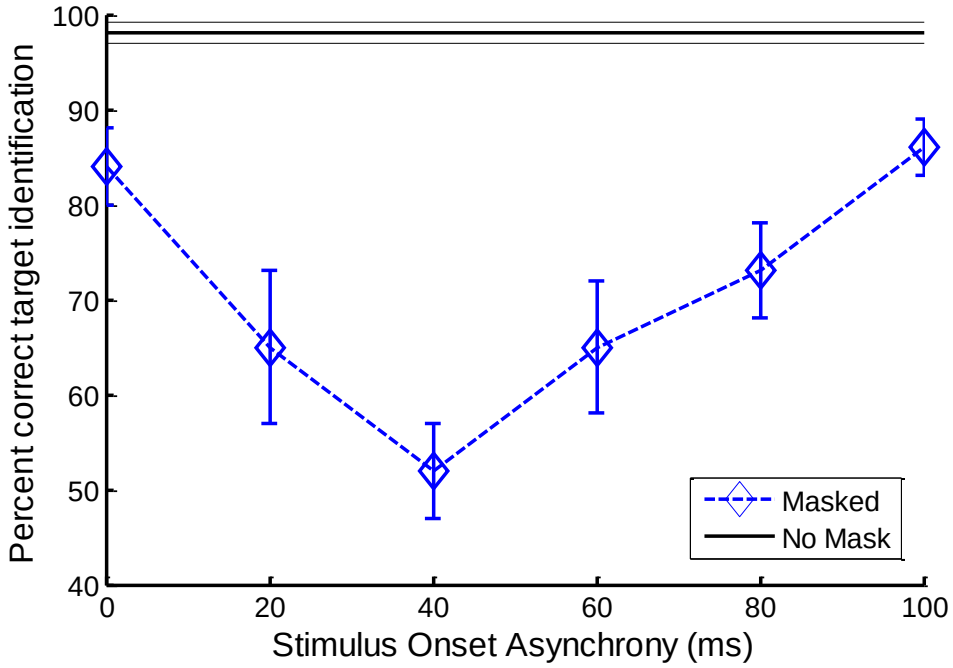


Figure 33: Static-Control Results from Experiment 1: Percent correct performance in detecting the missing corner of the target diamond (left/right), averaged across observers (N=4) and plotted against the SOA (ms). Subject performance is near chance at optimum SOA (40 ms). Error bars correspond to ±1 SEM. In the case of No Mask condition, ±1 SEM are shown by gray horizontal lines.

Figure 34 shows performance as a function of SOA for different experimental conditions, when Ternus-Pikler disks are perceived to be in element motion (ISI = 0ms). In comparison with the no mask condition, strong retinotopic (t_8 = -5.08; $p < 0.001$) but not non-retinotopic (t_8 = 1.57; p = 0.077) metacontrast masking is observed.

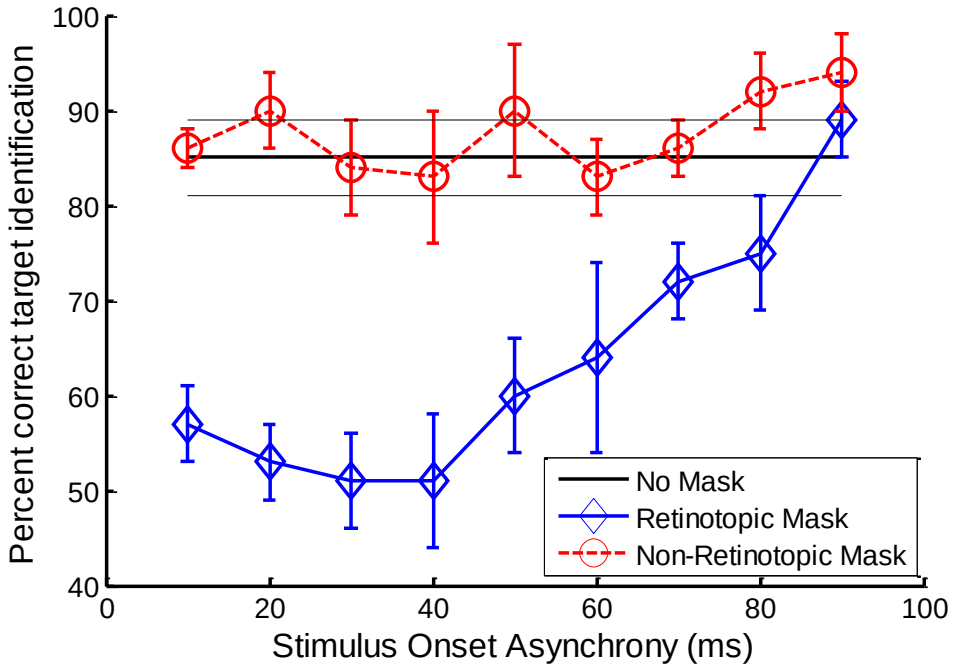


Figure 34: Element-Motion (ISI=0ms) Results from Experiment 1: Percent correct performance in detecting the missing corner of the target diamond (left/right), averaged across observers (N=4) and plotted against the SOA (ms), for ISI = 0ms. Subject performance is near chance at optimum SOA (40 ms) only in the retinotopic mask condition. Error bars correspond to ±1 SEM.

Figure 35 shows performance as a function of the SOA, when Ternus-Pikler disks are perceived to be in group motion (ISI = 40ms). In comparison with the no mask condition, significant retinotopic (t_4 = -3.46; p = 0.012) but not non-retinotopic (t_4 = 0.42; p = 0.347) masking is observed.

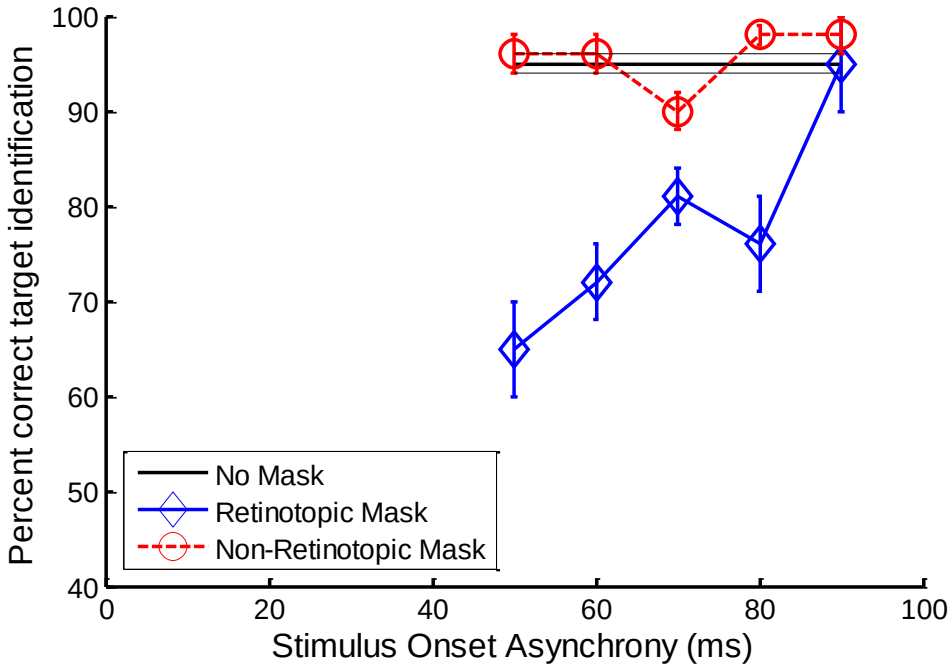


Figure 35: Group-Motion (ISI=40ms) Results from Experiment 1: Percent correct performance in detecting the missing corner of the target diamond (left/right), averaged across observers (N=4) and plotted against the SOA (ms), for ISI = 40ms. Error bars correspond to ±1 SEM. In the case of No Mask condition, ±1 SEM are shown by gray horizontal lines.

6.2.4 Discussion

These results indicate clearly that metacontrast masking in the absence of eye movements is retinotopic. Regardless of the perceived motion of Ternus-Pikler disks, the retinotopic mask significantly masks the target at optimum SOAs, while the presence of non-retinotopic mask has no significant effect on performance. In order to generalize this result across mask and masking function types, in the following experiment, we used a spatially *overlapping* mask that shared structural similarity with the target. In this structure masking paradigm, we chose a strong mask to generate a Type-A (i.e.

monotonic) backward masking function instead of the Type-B (i.e., non-monotonic) backward masking function obtained in the metacontrast experiment.

6.3 Experiment 8: Retinotopy of Masking by Structure

In order to investigate the retinotopy of structure masking in the absence of eye movement, the target/mask combination of Experiment 1 was modified as explained in the following section. Experimental procedures and results are discussed in the following sections.

6.3.1 Experimental Methods

Experimental design and procedures of Experiment 2 were identical to those of Experiment 1, with the exception of the target and mask design. The target consisted of a square outline missing one side. Three bars were aligned on the screen, as depicted in Figure 36-A, to form the target. The missing bar was randomly placed at the top or bottom of the square. The mask consisted of a collection of random horizontal and vertical bars, i.e. shared the same structural components as the target, to generate masking by structure (Figure 36-C). Figure 36-D displays the timing diagram of a typical trial in Experiment 2. Once again, the ISI was fixed (0 or 40ms) for each experimental block and the target always appeared just before the ISI. Mask presentation time, was again randomized from trial to trial to allow for different ISI-SOA combinations per block. Background luminance, Ternus-Pikler element shapes, and target/mask/disk contrasts were chosen as explained in Experiment 1. Ternus-Pikler radial motion (upward or downward) was also randomized from trial to trial. In a two-alternative forced-choice

design, three naïve observers and one of the authors reported the perceived location of the missing side of the target square (up/down).

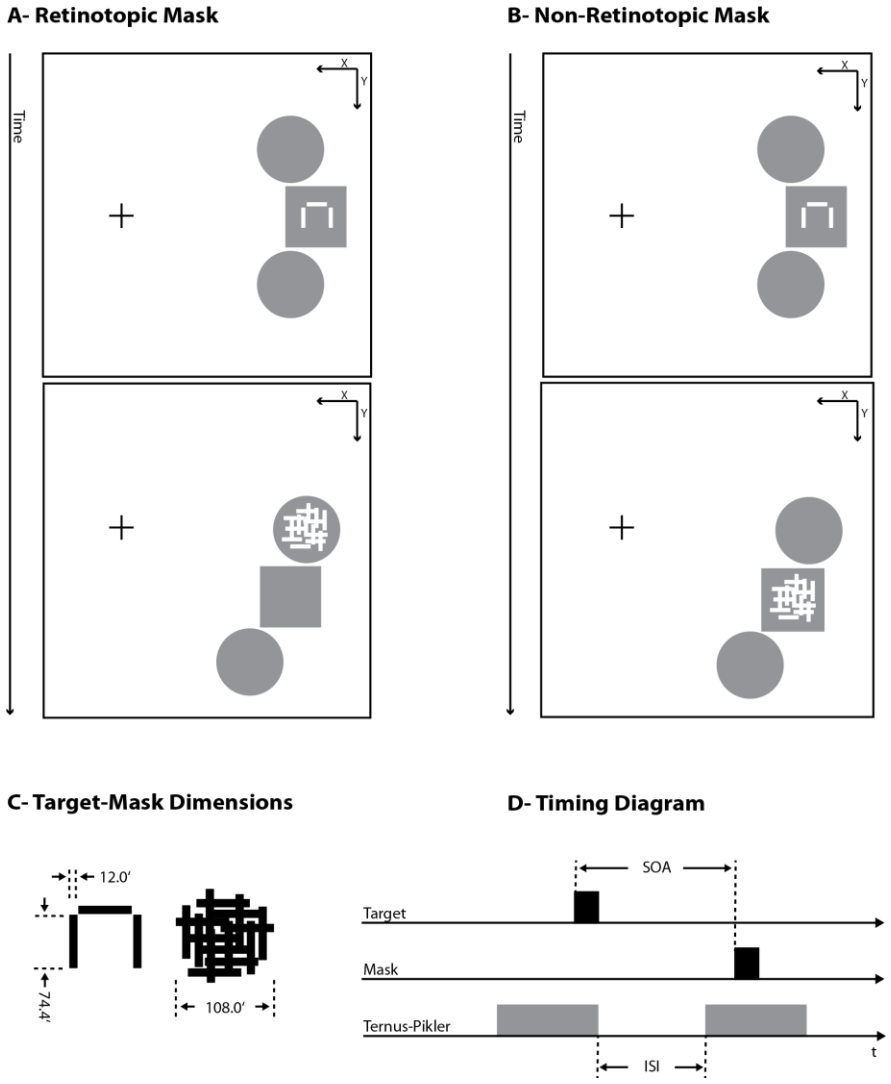


Figure 36: Stimuli and Respective Parameters for Experiment 2: (A) Retinotopic and (B) Non-retinotopic masking conditions. (C) The target consisted of a square outline missing one side. Three bars were aligned on the screen to form the target. The missing bar was randomly placed at the top or bottom of the square. The mask consisted of a collection of random horizontal and vertical bars. (D) Stimulus timing was identical to that of Experiment 1.

6.3.2 Results

Figure 37 shows performance as a function of the SOA for the static control condition.

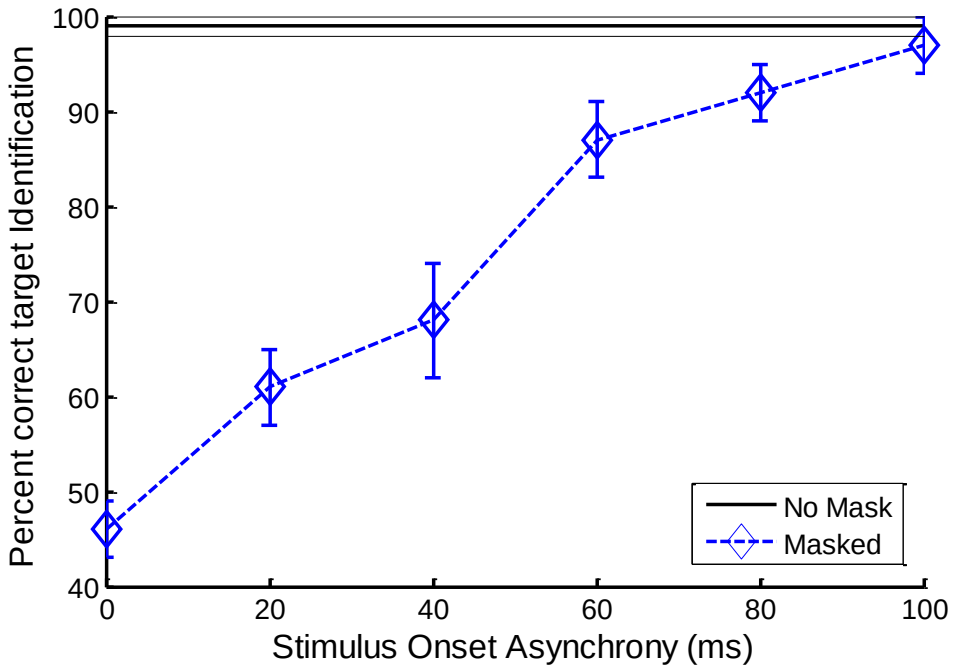


Figure 37: Static-Control Results from Experiment 2: Percent correct performance in detecting the missing side of the target square (up/down), averaged across observers (N=4) and plotted against the SOA (ms). Performance is near chance at optimum SOA (0 ms). Error bars correspond to ±1 SEM. In the case of No Mask condition, ±1 SEM are shown by gray horizontal lines.

Comparing to the no mask condition, the effect of the structure mask is significant (t_5 = 2.9; p = 0.016). As expected, performance results indicate a strong type A masking function. Figure 38 shows performance as a function of the SOA for the different experimental conditions, when Ternus-Pikler disks are perceived to be in element motion (ISI = 0ms). In comparison with the no mask experiment condition, retinotopic (t_4 = -3.98; p = 0.008) but not non-retinotopic (t_4 = -0.25; p = 0.404) masking effect is significant.

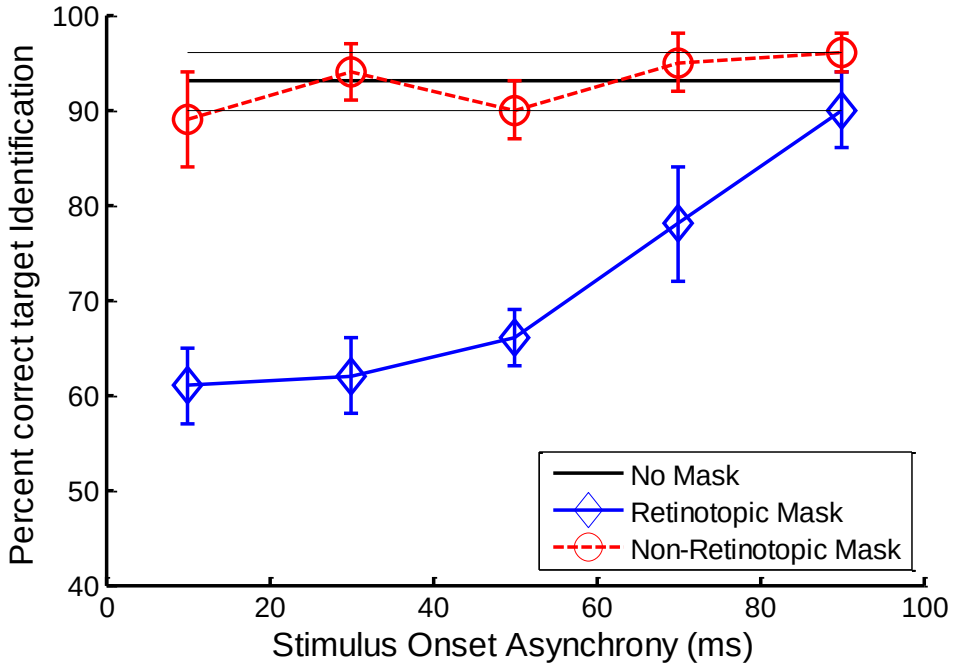


Figure 38: Element-Motion (ISI=0ms) Results from Experiment 2: Percent correct performance in detecting the missing side of the target square (up/down), averaged across observers (N=4) and plotted against the SOA (ms). Error bars correspond to ±1 SEM. In the case of No Mask condition, ±1 SEM are shown by gray horizontal lines.

Figure 39 shows performance as a function of the SOA for the different experimental conditions, when Ternus-Pikler disks are perceived to be in group motion (ISI = 40ms). In comparison with the no mask condition, significant retinotopic (t_4 = -2.94; p = 0.042) but not non-retinotopic (t_4 = -0.44; p = 0.341) masking effect is observed. Note that the pattern of results found under the Ternus-Pikler group motion condition (Figure 39) resembled those found under the element motion of the reference Ternus-Pikler disks (Figure 38).

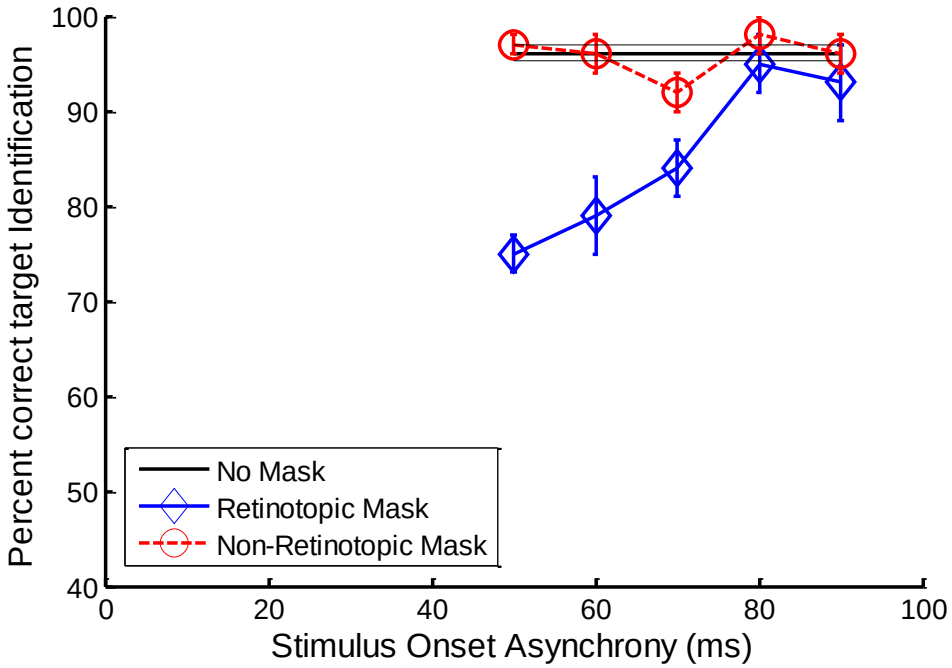


Figure 39: Group-Motion (ISI=40ms) Results from Experiment 2: Percent correct performance in detecting the missing side of the target square (up/down), averaged across observers (N=4) and plotted against the SOA (ms). Error bars correspond to ±1 SEM. In the case of No Mask condition, ±1 SEM are shown by gray horizontal lines.

6.3.3 Discussion

Taken together, the results of Experiments 1 and 2 show that backward masking is retinotopic and this finding holds for metacontrast and structure masking, as well as, for type-A and type-B masking functions. In a recent study, Lin and He (2012) investigated the retinotopy of masking by using a modified version of object-specific reviewing paradigm (Kahneman, Treisman, & Gibbs, 1992). A rectangular object (frame) was presented for a preview period of 200 ms. The target was presented during the last 10 ms of this preview period in one of the two sides of the rectangle. This rectangular frame was then shifted to a new location and displayed for another 200ms. The mask stimuli were presented during the first 30 ms of of the shifted frame. One side of the frame contained a

weak mask and the other side contained a strong mask. Neither mask occupied the same retinotopic location as the target but one of the masks occupied the same rectangle-relative position as the target (i.e., the same side). Observers performed better when the weak mask occupied the same relative position as the target. Lin and He interpreted this finding as evidence for non-retinotopic frame-centered backward masking. While this interpretation is plausible, it is difficult to make inferences about masking without observing the complete masking functions and comparing directly retinotopic, non-retinotopic, and baseline conditions. At the single short SOA of 10 ms (corresponding to ISI = 0 ms) used in the experiment, it is difficult to assess whether the difference in performance across the two mask types is due to masking *per se* or other factors. In our experiments, we included baseline no-mask measures, multiple SOA values to reveal the full typical type-A and type-B masking functions, and compared directly retinotopic and non-retinotopic masking conditions according to two different motion grouping conditions. Taken together, our results reveal only retinotopic masking.

Previous studies showed that features of a masked target can be observed as being part of the mask stimulus (Werner, 1935; Wilson & Johnson, 1985; Herzog & Koch, 2001; Otto et al., 2006; Ogmen et al., 2006; Breitmeyer et al, 2008). As indicated in Figure 30, the vernier offset of the target in the first frame can be observed on the mask stimulus shown in the second frame even though no vernier is presented at this element nor at its retinotopic location. Similarly, by using the sequential metacontrast paradigm, we have shown that features of a target, whose visibility is suppressed, can nevertheless be perceived along motion streams to which the target belongs (Otto et al., 2006; Herzog et al., 2012). Our previous studies have shown that the attribution of the target's features

to the mask stimulus is a consequence of motion grouping rather than masking itself (Ogmen et al., 2006; Breitmeyer et al., 2008). The goal of the next experiment was to study this motion-dependent non-retinotopic feature attribution in masking.

6.4 Experiment 9: Non-Retinotopic Feature Attribution under Visual Masking

In some trials of Experiments 7 and 8, subjects informally reported perceiving the target to be moving with the Ternus-Pikler elements, as one would expect from non-retinotopic feature attribution. In such cases, the target could be perceived at spatial locations different from where the target stimulus was actually presented. To formally study this phenomenon, we removed the motion ambiguity from Experiments 7 and 8, and instructed our subjects to spread their attention as discussed in the following section, so as to facilitate the read-out of non-retinotopic feature attribution.

6.4.1 Experimental Methods

Experimental design and procedures of Experiment 9 were identical to those of Experiments 7 and 8, with the exception of the Ternus-Pikler motion direction. In the previous experiments direction of the Ternus-Pikler radial motion was randomized clock-wise (downward) or counter-clockwise (upward) from trial to trial. The Ternus-Pikler motion in Experiment 9, however, was made predictably counter-clockwise (upwards) in all trials, and the subjects were instructed to spread their attention to *the central Ternus-Pikler square in both display frames*. The target and mask design was identical to those of Experiments 7 and 8 for the respective metacontrast and structure masking conditions. Stimulus timing was also chosen to match those of the previous two experiments. Once

again, the ISI was fixed (0 or 40ms) for each experimental block, and the target always appeared just before the ISI. Mask presentation time, was again randomized from trial to trial to allow for different ISI-SOA combinations per block. Background luminance, Ternus-Pikler element shapes, and target/mask/disk contrasts were chosen to match those of the previous experiments. The Ternus-Pikler radial motion was fixed (upward) in all trials to remove motion ambiguity. In a two-alternative forced-choice design, three naïve observers as well as one of the authors reported the perceived missing corner of the target diamond or the location of the missing side of the target square for metacontrast and structure masking conditions, respectively.

6.4.2 Results

Figures 40 shows performance as a function of the SOA for different structure mask experimental conditions, when Ternus-Pikler disks are perceived to be in element motion (ISI = 0 ms). In comparison with the no mask experiment condition, strong retinotopic ($t_8 = -3.82$; $p = 0.002$) but not non-retinotopic ($t_8 = -1.66$; $p = 0.067$) masking effect is observed. However, when the disks are perceived to be in group motion, masking effect becomes insignificant for both retinotopic ($t_4 = -1.12$; $p = 0.161$), as well as non retinotopic ($t_4 = 0.63$; $p = 0.280$) conditions (Figure 41). The significance of this finding will be discussed in detail in the next section, as well as the general discussion provided in chapter 8 of this dissertation.

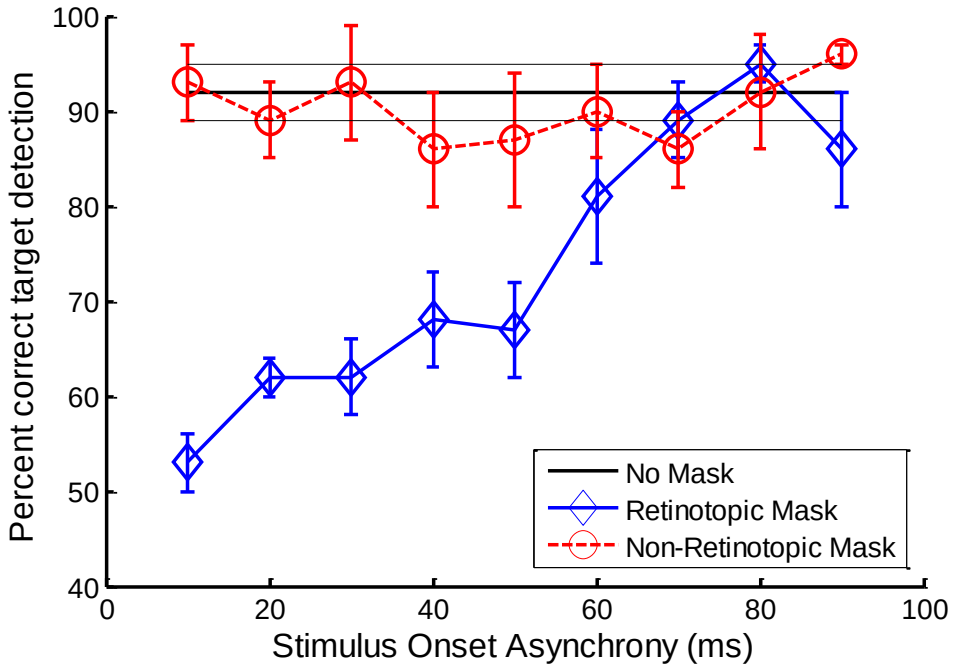


Figure 40: Structure-Mask Element-Motion (ISI=0ms) Results from Experiment 3: Percent correct performance in detecting the missing side of the target square (up/down), averaged across observers (N=4) and plotted against the SOA (ms), for ISI = 0ms. Subject performance is near chance at optimum SOA (10 ms) only in the retinotopic mask condition. Error bars correspond to ±1 SEM. In the case of No Mask condition, ±1 SEM are shown by gray horizontal lines.

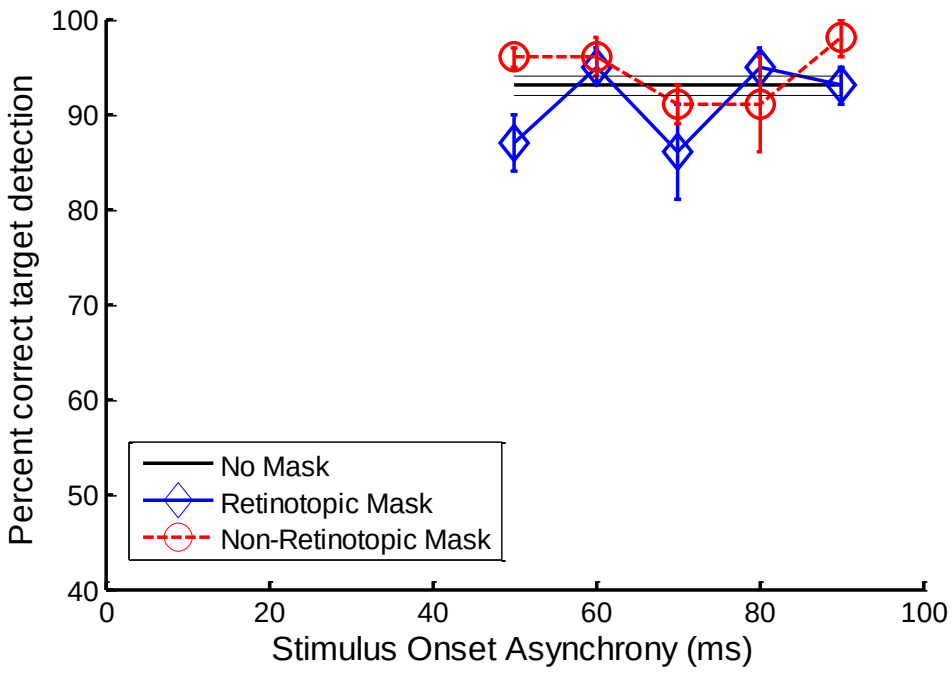


Figure 41: Structure-Mask Group-Motion (ISI=40ms) Results from Experiment 3: Percent correct performance in detecting the missing side of the target square (up/down), averaged across observers (N=4) and plotted against the SOA (ms). Error bars correspond to ±1 SEM. In the case of No Mask condition, ±1 SEM are shown by gray horizontal lines.

Similar pattern of results was observed in the case of metacontrast masking. Figure 42 shows performance as a function of the SOA, when Ternus-Pikler disks are perceived to be in element motion (ISI = 0ms). In comparison with the no mask experiment condition, strong retinotopic ($t_8 = -5.34$; $p < 0.001$) but not non-retinotopic ($t_8 = -1.69$; $p = 0.064$) masking effect is observed. However, when the disks are perceived to be in group motion, masking effect becomes insignificant for both retinotopic ($t_4 = -1.76$; $p = 0.076$) as well as non retinotopic ($t_4 = 0.30$; $p = 0.389$) conditions (Figure 41).

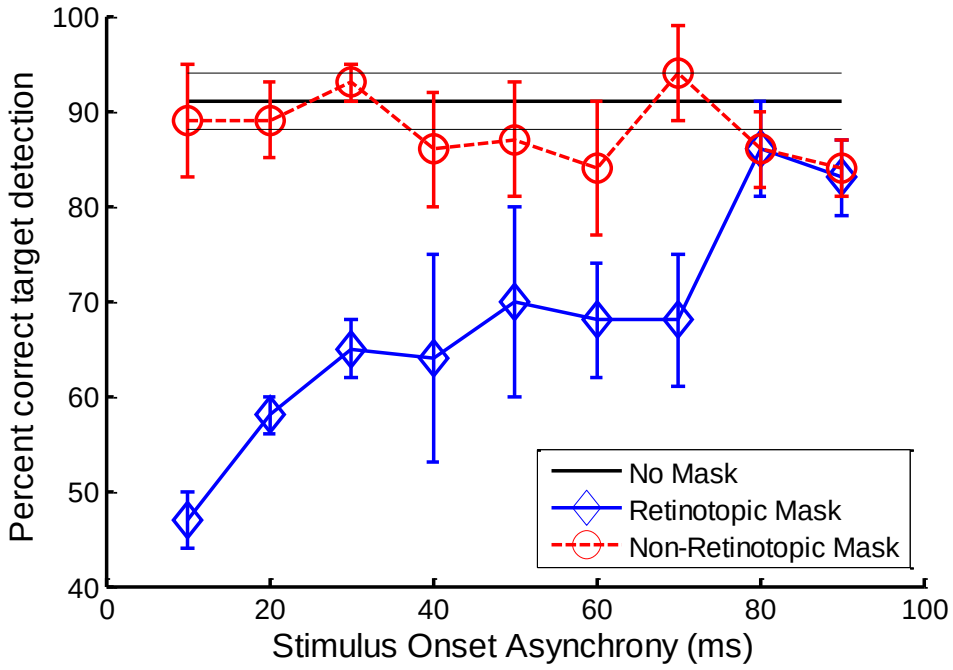


Figure 42: Metacontrast-Mask Element-Motion (ISI=0ms) Results from Experiment 3: Percent correct performance in detecting the missing corner of the target diamond (left/right), averaged across observers (N=4) and plotted against the SOA (ms). Subject performance is near chance at optimum SOA (10 ms) only in the retinotopic mask condition. Error bars correspond to ±1 SEM. In the case of No Mask condition, ±1 SEM are shown by gray horizontal lines.

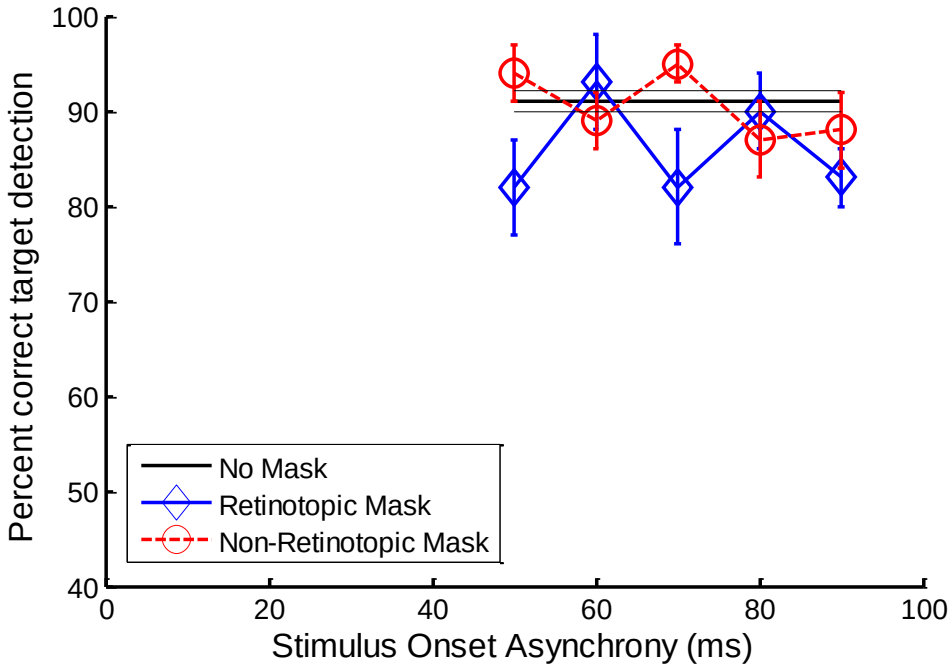


Figure 43: Metacontrast-Mask Group-Motion (ISI=40ms) Results from Experiment 3: Percent correct performance in detecting the missing corner of the target diamond (left/right), averaged across observers (N=4) and plotted against the SOA (ms). Error bars correspond to ±1 SEM. In the case of No Mask condition, ±1 SEM are shown by gray horizontal lines.

6.4.3 Discussion

In agreement with the results found in experiments 7 and 8, retinotopic masking is observed when the Ternus-Pikler disks are perceived in element motion (ISI = 0ms). However, in contrast to the results found in our previous two experiments, when observers can focus their attention to the Ternus-Pikler element in the second frame which is grouped with the Ternus-Pikler element in the first frame containing the target, they can identify the target based on its continued appearance along the motion path of the element containing the target. This finding is in agreement with our previous results from sequential metacontrast (Otto et al., 2006; Herzog et al., 2012) and Ternus-Pikler

display (Figure 30). Informal reports of our subjects state that a faded, but complete copy of the target is perceived at the non-retinotopic destination, in accordance with the motion of the stimulus.

6.5 Eye-Movement Controlled Experiments

Our previous experiments indicate that observers are able to keep a stable fixation while viewing the Ternus-Pikler displays (Boi et al., 2011). Nevertheless, to completely rule out the involvement of eye movements, we repeated experiment 9 in a control experiment with eye movement monitoring.

6.5.1 Methods

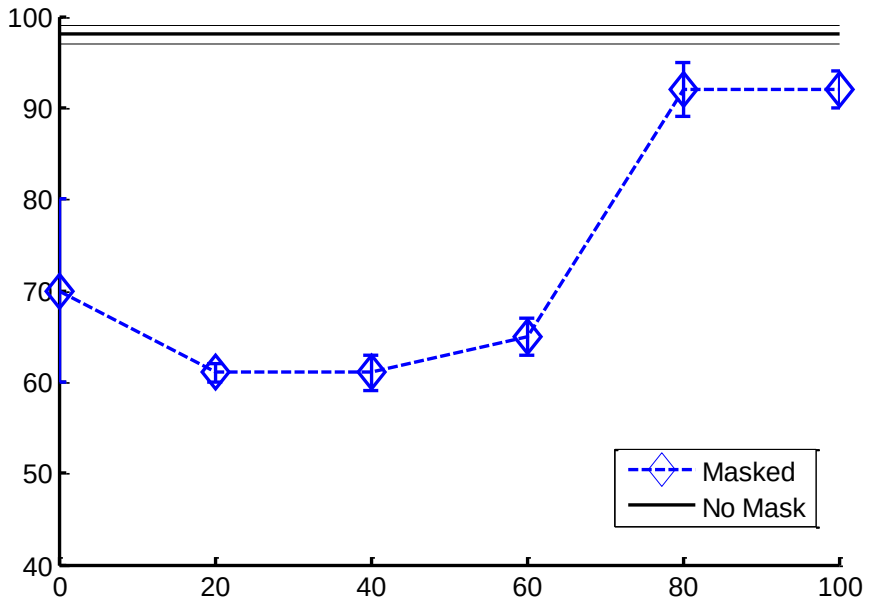
The experimental procedures and stimulus parameters for the eye movement control experiments were identical to those of the original experiment, with the exception of eye monitoring procedures as explained below. Three observers (one new and two from the original participants) were stablilized using a chin rest, and eye position was sampled at a rate of 250 Hz using an SR-Research Eyelink II eye tracker with default saccade detection. A nine-point calibration was conducted at the beginning of each experiment block to map observer eye position to screen coordinates. Drift correct routine was conducted before every trial to account for minor observer or headgear movement during each block. Observer eye movements were analyzed online and offline, throughout all trials. Trials during which a saccade was detected were rejected and repeated as a new trial. Similarly, all trials during which observer gaze moved outside an imaginary circle of 1° diameter (centered on the fixation point) were rejected and repeated online. Furthermore, offline analysis was conducted on the eye movement data

to rule out any correlation between eye movement and performance. Analysis of eye movement data indicates that subjects were able to maintain highly accurate fixation throughout the trials.

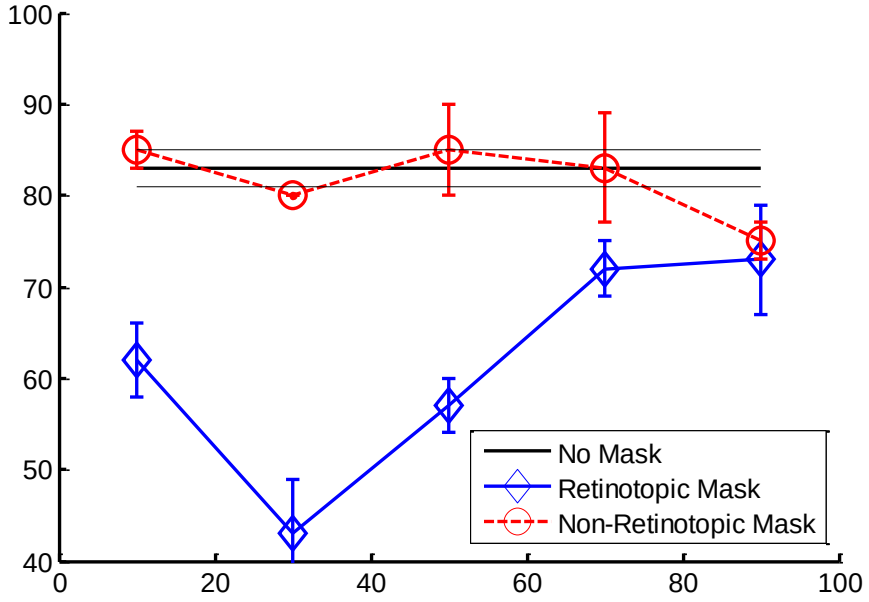
6.5.2 Results

In general, the results of eye movement control experiments were in agreement with earlier findings. Stimulus parameters were chosen to achieve strong metacontrast masking in the absence of Ternus-Pikler motion, as depicted in the plot of performance as a function of the SOA in Figure 44-A (t_5=4.069; p<0.001; d=1.57).

A – Static Control



B- Element Motion (ISI = 0 ms)



C- Group Motion (ISI = 40 ms)

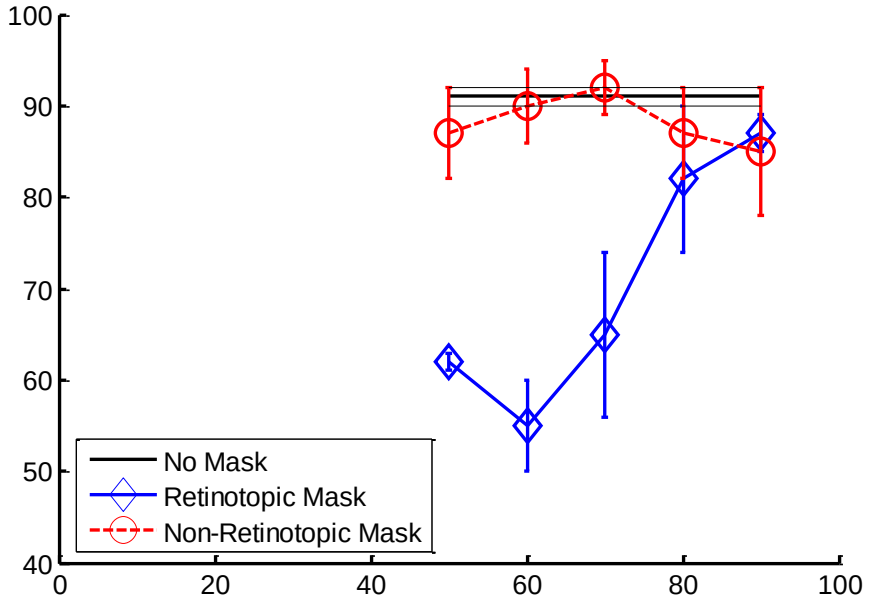
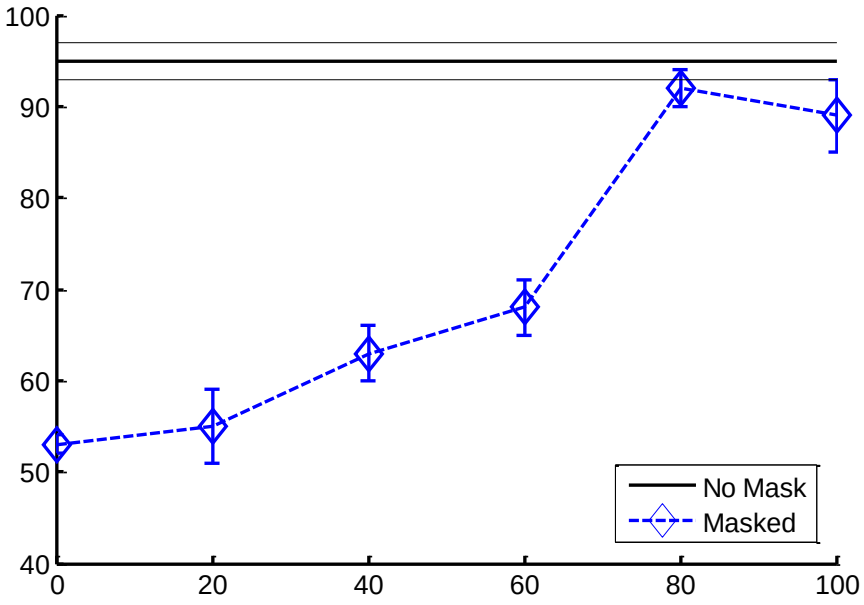


Figure 44: Metacontrast masking. Percentage of correct responses in detecting the missing corner of the target diamond (left/right), averaged across observers (N=4). A. Static control condition. Performance is near chance at an SOA of 40 ms with a Type-B masking function. B. Element-Motion (ISI=0ms). Performance

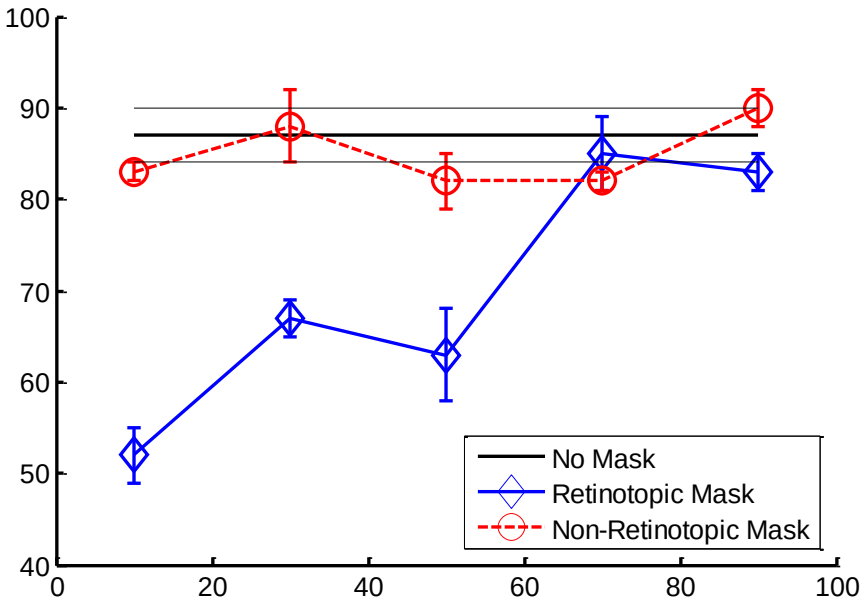
is near chance at SOAs near 40 ms only in the retinotopic mask condition. C. Group-Motion (ISI=40ms). Masking is observed only for the retinotopic mask. Error bars correspond to ±1 SEM. In the case of No Mask condition, ±1 SEM are shown by gray horizontal lines.

Under Ternus-Pikler element motion of the disks, in comparison with the no mask condition, strong retinotopic (t_4=3.966; p=0.016; d=1.62) but no non-retinotopic (t_4=0.701; p=0.52) metacontrast masking is observed (Figure 44-B). When Ternus-Pikler disks are perceived to be in group motion (ISI = 40ms), in comparison with the no mask condition, significant retinotopic (t_4=3.468; p=0.025; d=1.54) but not non-retinotopic (t_4=2.449; p=0.070) masking is observed (Figure 44-C).

A – Static Control



B- Element Motion (ISI = 0 ms)



C- Group Motion (ISI = 40 ms)

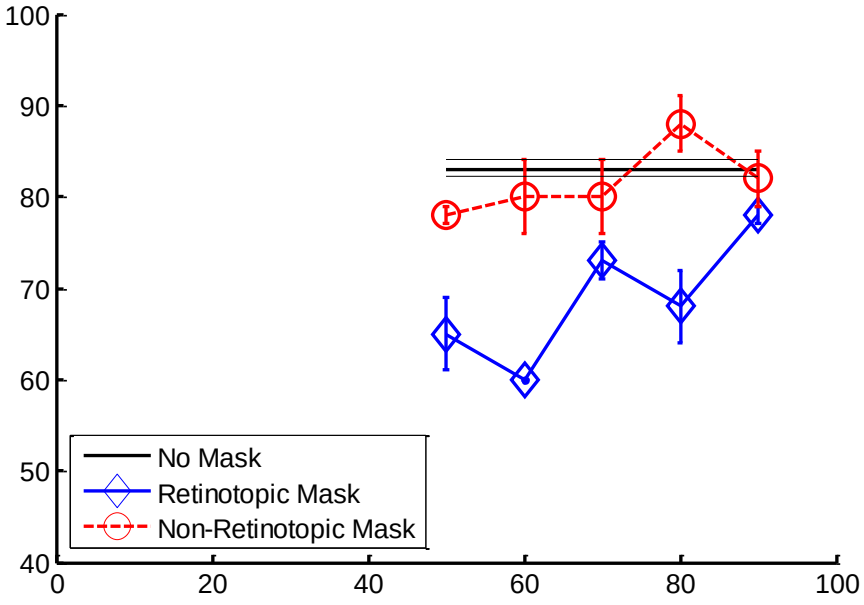
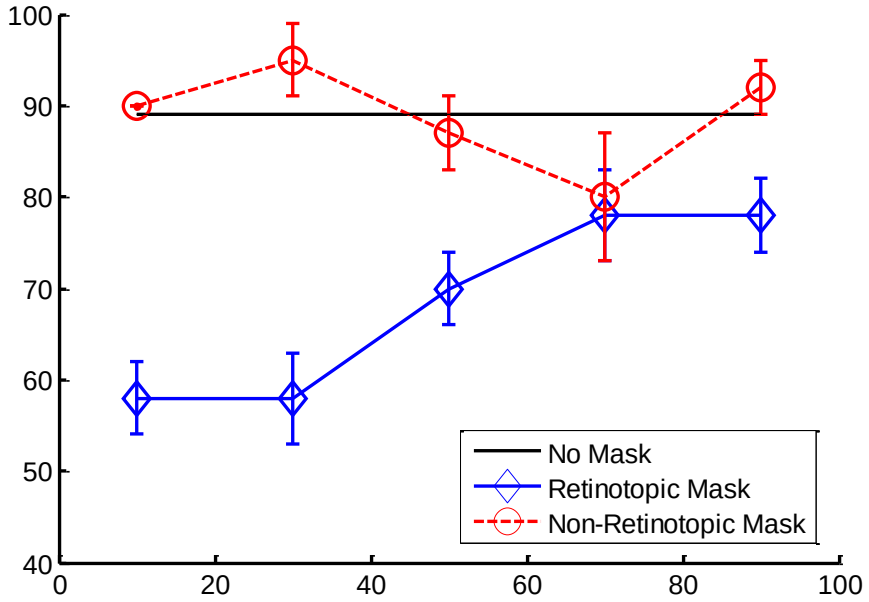


Figure 45: Masking by structure. Percentage of correct responses in detecting the missing side of the target square (up/down), averaged across observers (N=4).. A. Static control condition. Performance is near chance at SOA of 0 ms with a Type-A

masking function. B. Element-Motion (ISI=0ms). C. Group-Motion (ISI=40ms). Error bars correspond to ±1 SEM. In the case of No Mask condition, ±1 SEM are shown by gray horizontal lines.

Similar masking effects were observed when a structure mask was utilized. Figure 45-A shows performance as a function of the SOA for the static control condition (in the absence of Ternus-Pikler motion). Compared to the no mask condition, significant type-A masking was found (t_5=2.921; p=0.032; d=1.35). When Ternus-Pikler disks are in element motion (ISI = 0ms), in comparison with the no mask condition, significant retinotopic (t_4=-3.986; p=0.008; d=1.63) but not non-retinotopic (t_4=-0.256; p=0.809) masking effect was observed (Figure 45-B). Once again, similar masking effects were observed when Ternus-Pikler disks were in group motion (ISI = 40ms). In comparison with the no mask condition, significant retinotopic (t_4=-2.942; p<0.042; d=1.44) but not dnon-retinotopic (t_4=-0.44; p=0.682) masking effect was observed (Figure 45-C).

A- Element Motion (ISI = 0 ms)



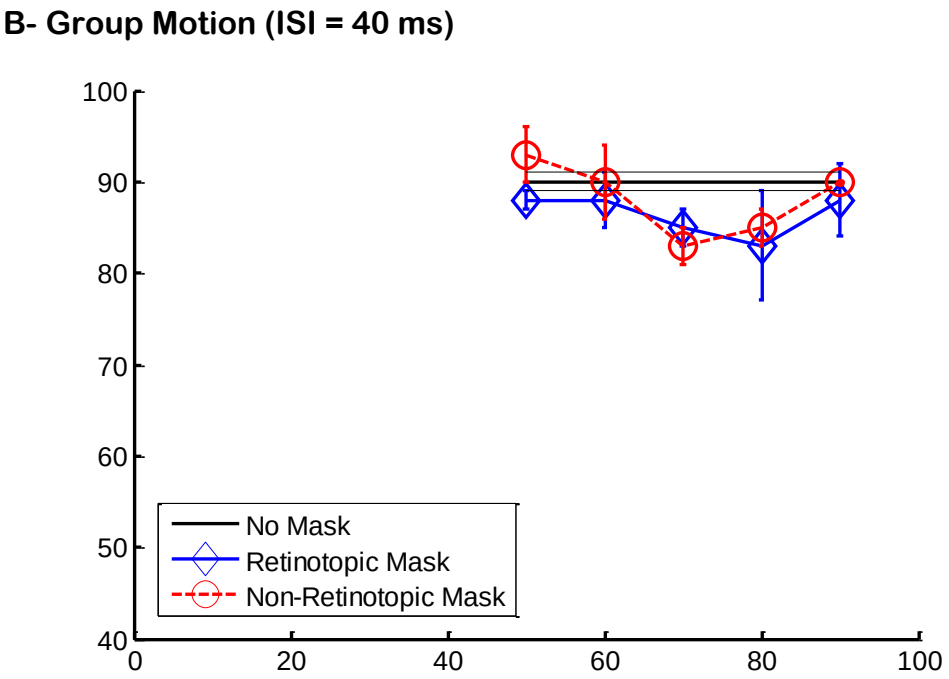
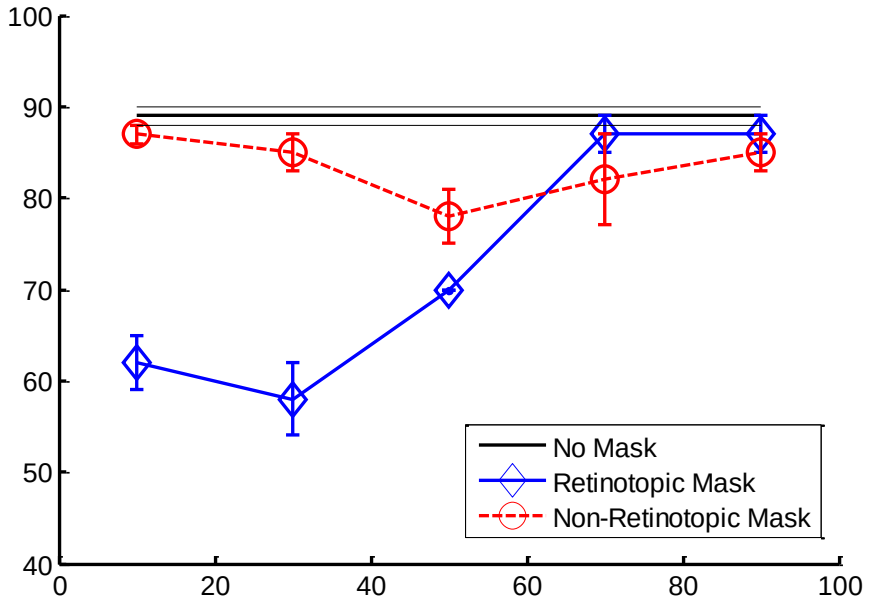


Figure 46: Metacontrast masking with predictable Ternus-Pikler motion. The observers attended to the central Ternus-Pikler square in both display frames. Percentage of correct responses in detecting the missing side of the target diamond (left/right), averaged across the four observers. A. Element-Motion (ISI=0ms). Performance is near chance at SOA of 10 ms only in the retinotopic mask condition. B. Group-Motion (ISI=40ms). No masking is observed. Error bars correspond to ±1 SEM. In the case of No Mask condition, ±1 SEM are shown by gray horizontal lines.

Furthermore, under predictable Ternus-Pikler reference motion, the results of eye movement control experiments remain in agreement with the findings reported in Experiment 3. Figure 46-A shows performance as a function of the SOA, when Ternus-Pikler disks are perceived to be in predictable element motion (ISI = 0ms). In comparison with the no mask experiment condition, significant retinotopic (t_4=4.459; p=0.011; d=1.68) but not non-retinotopic (t_4=0; p=1) metacontrast masking effect is observed. However, when the disks are perceived to be in predictable group motion (Figure 46-B),

masking effect becomes insignificant for both retinotopic (t_4=2.213; p=0.091) as well as non retinotopic (t_4=0.365; p=0.733) conditions.

A- Element Motion (ISI = 0 ms)



B- Predictable Structure (ISI = 40)

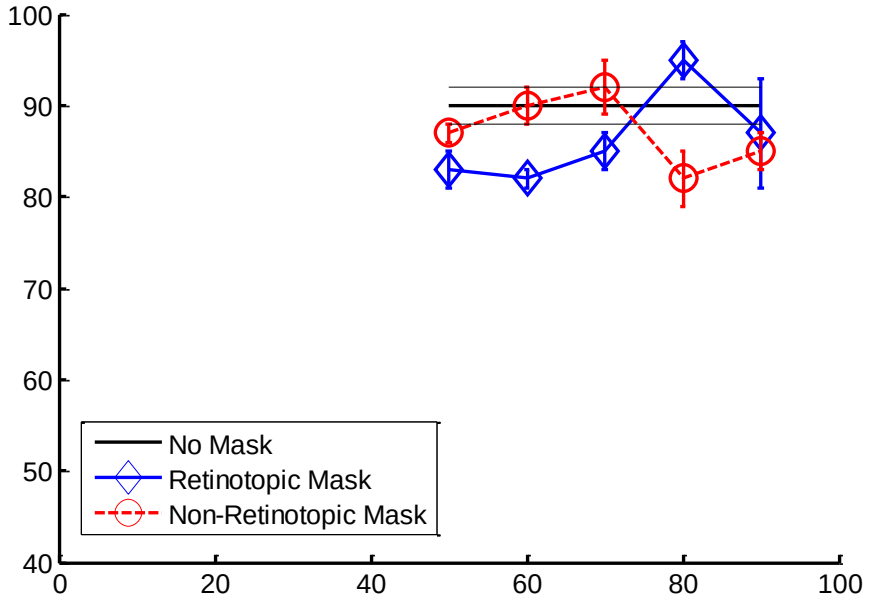


Figure 47: Masking by structure with predictable Ternus-Pikler motion. The observers attended to the central Ternus-Pikler square in both display frames. Percentage of correct responses in detecting the missing side of the target square (up/down), averaged across the four observers. A. Element-Motion (ISI=0ms). Performance is near chance at SOA of 10 ms only in the retinotopic mask condition. B. Group-Motion (ISI=40ms). No masking is observed. Error bars correspond to ±1 SEM. In the case of No Mask condition, ±1 SEM are shown by gray horizontal lines.

Similar effects were observed in the masking by structure experiments, with only one exception. As depicted in Figure 47-A, under predictable Ternus-Pikler element motion (ISI=0ms), masking effect was found significant for retinotopic (t_4=2.71; p=0.05; d=1.38) as well as non-retinotopic (t_4=3.80; p=0.019; d=1.6) mask conditions. This observation was rather surprising as no evidence was found for non-retinotopic masking effect in any of our earlier experiments. A closer look at these results, however, reveals that in the case of the non-retinotopic mask condition, subject performance remains near 80% at optimal SOA (i.e. 50 ms) and near 90% at shorter SOAs (e.g. 10 ms). High performance levels and the type-B shape of the masking function collectively suggest that the observed effect cannot be attributed to masking by structure. Nonetheless, under predictable Ternus-Pikler group motion, both retinopic (t_4=1.548; p=0.196) and non-retinotopic (t_4=1.685; p=0.167) masking effects cease to be significant (Figure 47-B). These results collectively support our earlier findings and rule out eye movement as an explanation for elimination of masking effect under predictable group motion condition.

Chapter 7 Future Work

The essential role of motion segmentation in formation and maintenance of non-retinotopic reference frames suggests significant interaction between the primary visual pathways for computation of dynamic form. The experiments discussed below can investigate involvement of the cortical regions involved in formation and maintenance of non-retinotopic reference frame.

7.1 Background

Despite recent advancements, neurophysiological knowledge of primate brain remains coarse-grained and cannot be directly mapped to our theory. The early visual areas V1, V2, V3, V4/8 and V3a have been extensively researched and are known to be retinotopic. Virtually, any normal visual stimulation activates V1 and V2, in humans and other primates alike. Beyond the retinotopic areas, however, neurophysiological findings become less pronounced. Yin et al., investigated the neural correlates of a non-retinotopic visual phenomenon called anarthoscopic perception using Functional Magnetic Resonance Imaging (fMRI) techniques (Yin, Shimojo, Moore, & Engel, 2002). In the absence of correlation between retinotopic activities and their subjects' perception of the anorthoscopic stimulus, Yin et al. reported cortical activities at non-retinotopic areas such as the Lateral Occipital Complex (LOC) and along the motion pathway in M+ area. Their results hint that the LOC and other non-retinotopic areas could be potential candidates for creation and maintenance of non-retinotopic reference frames.

The experiments discussed in this chapter, can allow us to investigate neural correlates of non-retinotopic representations using fMRI techniques. These proposed experiments are designed to be conducted following an fMRI retinotopic mapping (Engel, 1994) of the visual cortex for every subject.

7.2 Experiment F1: fMRI Map of Ternus-Pikler in Element and Group Motion

The basic ternus-pikler stimulus described in the earlier chapters can be modified to include the flickering checkerboard pattern as shown in Figure 48.

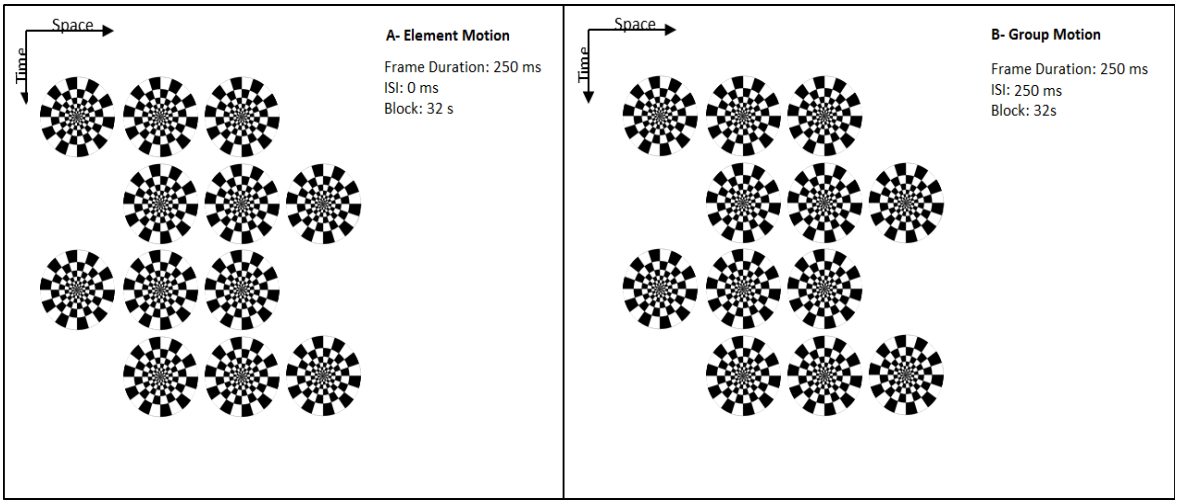


Figure 48: Stimuli for Future Experiment I: Ternus-Pikler Element (A) and Group (B) Motion Condition

In the proposed experiment, fMRI scans are to be recorded during stimulus presentation for both element and group motion percepts. The obtained fMRI maps can then be processed and analyzed to investigate non-retinotopic areas associated with the

perception of group motion. Contrasting fMRI maps of element motion and group motion conditions can shed light on neural correlates of Ternus-Pikler phenomenon.

7.3 Experiment F2: Neural Correlates of Non-Retinotopic Activities

A Ternus-Pikler experimental design similar to that of Figure 10, though replacing the target dot with a flickering wedge, can provide a suitable paradigm for exploring the neural correlates of non-retinotopic reference frames. Subjects will keep fixation at the center of the screen while attending to the wedge motion in order to perceive direction of wedge rotation.

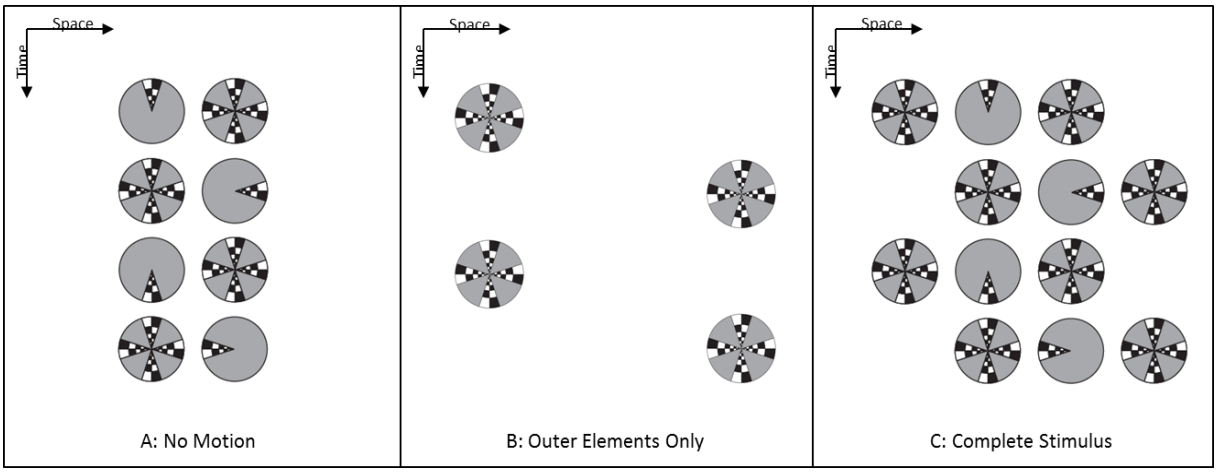


Figure 49: Stimuli for Future Experiment II: (A) No Motion Condition. (B) Flanker Conditions. (C) Full Ternus-Pikler Group Motion Condition

The stimulus design includes three conditions as shown in Figure 49: A) no-motion condition; B) outer elements only condition; and C) the complete Ternus-Pikler stimulus. We hypothesize that contrasting the full Ternus-Pikler stimulus (C) and the sum of the other two cases (B+C) in fMRI maps can highlight the neural regions selective to non-retinotopic activities that lead to the perception of wedge rotation under group

motion condition. Moreover, applying retinotopic mapping techniques tailored for rotating wedges (Engel, 1994) to the regions of interest (ROIs) identified by the contrast mentioned above, can determine any dynamic neural map activated by the reference frame.

Chapter 8 General Conclusions and Discussion

Our results indicate that the effect of a dynamic reference frame extends over space and time, creating a field within which target stimuli are localized and perceived relative to the reference frame. The strength of the field effect is independent of the inducing element size. The fields of neighboring dynamic reference frames interact, while static references do not affect the fields of neighboring dynamic references. The magnitude of this interaction increase as the distance between the two neighboring references decreases. Temporal synchronization also plays a significant role with respect to the field effect. The magnitude of the non-retinotopic field effect on the target stimulus is maximized when the target and the reference stimuli appear in phase. The field strength decreases as the target-reference temporal asynchrony is increased. These results collectively emphasize that the reference frame revealed by our studies can be better described in terms of a "field" rather than an object. Our results also indicate that inter-reference interactions occur only when these neighboring references are in motion; suggesting that the fields generated by non-retinotopic reference frames are motion-based. The results of our visual masking experiments indicate that while visual masking mechanisms operate in retinotopic domain, masking effect attenuates significantly in the presence of predictable non-retinotopic reference frames. This finding is consistent with the theory of dynamic form perception (Ogmen, 2007), where features are mapped to non-retinotopic loci, based on the reference frame induced by the stimulus.

The nervous system uses a variety of reference frames according to different tasks it performs. For example, a body-centered reference frame is especially useful in coordinating the interactions of the body and the limbs with the environment. A body

centered reference frame can guide reaching movements since the variable of interest is the position of the selected target with respect to the hand. A retinotopic reference frame can effectively produce an error signal to move or to keep the fovea on a select target. These types of reference frames, which are relative to the observer, are called egocentric (viewer-centered) reference frames. Coordination between different senses or between perception and action require coordination between their respective reference frames. In early stages of cognitive development, the child's universe is built mainly around egocentric reference frames; however, later in development, the child undergoes a "decentering" process whereby exocentric (also known as allo-centric, or non-viewer-centered) reference frames lead to an appreciation of a world independent of the self (Montangero & Maurice-Neville, 1994; Piaget & Inhelder, 1969). Exocentric reference frames are those that are relative to entities outside of the observer. Since in our study the observer is stationary with respect to the stimulus, the non-retinotopic effects that we observe can be attributed to an exocentric reference frame. Exocentric reference frames play a significant role in computations that determine observer-independent properties of stimuli, such as view-point invariant recognition of objects. Two commonly evoked exocentric reference frames are spatiotopic and object-based reference frames. The former refers to a reference frame fixed at a given location in space and thus remains stationary with respect to the space surrounding the observer. The latter refers to a reference frame fixed on an object. When the object is stationary, object-based and spatiotopic reference frames become equivalent. However, when the object moves, the reference frame is no longer stationary in space but moves with the object. Since in our experiments observers remained stationary with respect to the environment, the effects

that we observe cannot be explained by spatiotopic reference frames. Does an object-based reference frame constitute an appropriate way to describe our findings? The term "object", although intuitively appealing, is rather vague in its definition (Avrahami, 1999; Egly, Driver, & Rafal, 1994; Feldman, 1999; Humphreys & Riddoch, 2007; Kasai, Moriya, & Hirano, 2011; Marino & Scholl, 2005; Marr, 1982; Scholl, Pylyshyn, & Feldman, 2001). Considering the commonly suggested constraints of closure and connectedness to define objects, we suggest that the reference frame revealed by our studies can be better described in terms of a "field" rather than an object. We use the term field in a similar way to its use in Gestalt psychology, which in turn is an adaptation of the field concept from physics (Koffka, 1935). From a more modern perspective, the field effect can be expressed as curvature of perceptual space-time (cf., gravity lens theory for a static version: (Greene, 1998; Naito & Cole, 1994). The traditional definition of object (closure and connectivity) would suggest interactions via a direct physical mediator (e.g., movement of the torso inducing the movement of the limb) while the field concept allows to explain how effects can spread over space without requiring physical contact or connectivity.

In the work presented in this dissertation, by using motion perception as an example, we examined how an exocentric reference frame exerts its influence on stimuli. Our results show that the effect of the reference frame is independent of the inducing element size, and spreads over space uniformly. Whether the probe stimuli (dots) were placed inside or outside of the disks had no effect. Similarly, when a second static reference frame was introduced (Experiment 3, 4), whether or not the probe dot fell inside the elements of this second reference frame had no influence on our results. Thus placing the

dot inside the putative reference object and making it part of that object versus placing it outside without a connection had no effect. Similarly, placing the dot inside another object (the neighboring square) had no effect either. The proximity of the dot to the object had no effect within the tested range. Thus the effect of the reference frame spreads uniformly (within the tested range) over space as a field, influencing other stimuli presented within this field. Our results also indicate that the interactions between reference frames occur only when they are in motion; suggesting that the fields generated by the reference frames are motion-based. Previous work, using the Ternus-Pikler displays, have shown that form (Ogmen et al., 2006), visual search (Boi et al., 2009), and exogenous attention (Verger et al., 2011) occur according to a dynamic non-retinotopic reference frame. Future research will determine whether field effects apply to these processes as found in the case of motion perception.

The functional significance of retinotopic masking in the absence of eye movement can be understood by considering how the visual system analyzes the form of moving targets. Under normal viewing conditions, a briefly presented stimulus remains visible for approximately 120 ms after the stimulus offset (Coltheart, 1980; Haber & Standing, 1970). Due to this visible persistence, one would expect moving objects to appear highly blurred with a comet-like trailing smear. Yet our normal perception of objects in motion is relatively clear and sharp (Bex et al., 1995; Hammett, 1997; Ramachandran et al., 1974), a phenomenon known as motion deblurring (Burr, 1980; Burr & Morgan, 1997; Chen et al., 1995; Dilollo & Hogben, 1985). It has been suggested that motion deblurring results from visual masking mechanisms (Ogmen, 1993; Chen, et al. 1995; Purushothaman et al., 1998; Ogmen & Breitmeyer, 2006; Ogmen, 2007). According to

this account, nearby stimuli suppress the visibility of motion streaks generated by moving targets (Breitmeyer & Ogmen, 2006; Chen et al., 1995; Ogmen, 1993, 2007; Purushothaman et al., 1998). However, masking mechanisms solve only partly the motion blur problem. They can make motion streaks appear shorter thereby reducing the amount of blur in the picture. Yet, although deblurred, moving objects would still suffer from having a ghost-like appearance (Ogmen, 2007). This is because in the retinotopic space, a moving object will stimulate each retinotopically localized receptive-field briefly. Insufficient stimulation leads to spread of incompletely processed form information across the retinotopic space, just like the ghost-like appearances of moving objects in pictures taken at relatively slow shutter speeds. As a solution to this "moving ghosts" problem, our findings suggest that features of moving objects are processed according to motion-based non-retinotopic reference frames. Future studies will determine how reference frames generated by ego-motions (as in the case of eye movements) and exo-motions (as in the case of moving objects, studied herein) are coordinated to work in synergy. The experiments presented in this work, however, collectively support the hypothesis that clarity of vision is achieved through a synergy between grouping, retinotopic masking, and non-retinotopic feature attribution.

References

Alpern, M. (1952). Metacontrast; Historical Introduction. *American journal of optometry and archives of American Academy of Optometry, 29*(12).

Anderson, C. H., & Van Essen, D. C. (1987). Shifter Circuits - A Computational Strategy for Dynamic Aspects of Visual Processing. *Proceedings of the National Academy of Sciences of the United States of America, 84*(17).

Anderson, C. H., & Vanessen, D. C. (1987). Shifter Circuits - A Computational Strategy for Dynamic Aspects of Visual Processing. *Proceedings of the National Academy of Sciences of the United States of America, 84*(17).

Avrahami, J. (1999). Objects of Attention, Objects of Perception. *Perception & Psychophysics, 61*(8), 1604-1612.

Aydin, M., Herzog, M. H., & Ogmen, H. (2011). Attention Modulates Spatio-Temporal Grouping. *Vision Research, 51*(4), 435-446.

Bachmann, T. (1994). Psychophysiology of Visual Masking: The Fine Structure of Conscious Experience. Commak, NY: Nova Science.

Bedell, H. E., & Lott, L. A. (1996). Suppression of Motion-Produced Smear During Smooth Pursuit Eye Movements. *Current Biology, 6*(8).

Bertrand, O., & Tallon-Baudry, C. (2000). Oscillatory Gamma Activity in Humans: A Possible Role for Object Representation. *International Journal of Psychophysiology, 38*(3), 211-223.

Bex, P. J., Edgar, G. K., & Smith, A. T. (1995). Sharpening of Drifting, Blurred Images. *Vision Research, 35*(18).

Bidwell, S. (1899). Curiosities of Light and Sight. London: Swan Sonnenschein.

Biederman, I. (1987). Recognition-by-Components - A Theory of Human Image Understanding. *Psychological Review, 94*(2).

Biederman, I., & Gerhardstein, P. C. (1993). Recognizing Depth-Rotated Objects - Evidence and Conditions for 3-Dimensional Viewpoint Invariance. *Journal of Experimental Psychology-Human Perception and Performance, 19*(6), 1162-1182.

Boi, M., Ogmen, H., & Herzog, M. H. (2011). Motion and Tilt Aftereffects Occur Largely in Retinal, not in Object, Coordinates in the Ternus-Pikler Display. *Journal of vision, 11*(3).

Boi, M., Ogmen, H., Krummenacher, J., Otto, T. U., & Herzog, M. H. (2009). A (Fascinating) Litmus Test for Human Retino- vs. Non-Retinotopic Processing. *Journal of Vision, 9*(13), 11.

Braddick, O. (1973). Masking of Apparent Motion in Random-Dot Patterns. *Vision Research, 13*(2), 355-369.

Braddick, O. (1974). Short-Range Process in Apparent Motion. *Vision Research, 14*(7), 519-527.

Braddick, O., & Adlard, A. (1978). *Apparent Motion and the Motion Detector* New York, NY Academic Press,.

Breitmeyer, B. G., & Ogmen, H. (2006). Visual Masking: Time Slices through Conscious and Unconscious Vision (2nd Edition ed.). Oxford, UK.: Oxford University Press.

Breitmeyer, B. G., & Ritter, A. (1986a). The Role of Visual-Pattern Persistence in Bistable Stroboscopic Motion. *Vision Research, 26*(11), 1801-1806.

Breitmeyer, B. G., & Ritter, A. (1986b). Visual Persistence and the Effect of Eccentric Viewing, Element Size, and Frame Duration on Bistable Stroboscopic Motion Percepts. *Perception & Psychophysics, 39*(4), 275-280.

Burr, D. (1980). Motion Smear. *Nature, 284*(5752).

Burr, D., & Morgan, M. J. (1997). Motion deblurring in human vision. *Proceedings of the Royal Society B-Biological Sciences, 264*(1380).

Castet, E. (1994). Effect of the ISI on the Visible Persistence of a Stimulus in Apparent Motion.

Cavanagh, P., & Mather, G. (1989). Motion: the long and short of it. *Spatial vision, 4*(2-3), 103-129.

Chen, S., Bedell, H. E., & Ogmen, H. (1995). A Target in Real Motion Appears Blurred in the Absence of Other Proximal Moving Targets. *Vision Research, 35*(16).

Coltheart, M. (1980). The Persistencies of Vision. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences, 290*(1038).

Cutting, J. E., & Proffitt, D. R. (1982). The Minimum Principle and the Perception of Absolute, Common, and Relative Motion. *Cognitive Psychology, 14*(2).

Davidson, M. L., Fox, M. J., & Dick, A. O. (1973). Effect of Eye-Movements on Backward Masking and Perceived Location. *Perception & Psychophysics, 14*(1), 110-116.

Dilollo, V., & Hogben, J. H. (1985). Suppression of Visible Persistence. *Journal of Experimental Psychology-Human Perception and Performance, 11*(3).

Dixon, N. F., & Hammond, E. J. (1972). Attenuation of Visible Persistence. *British Journal of Psychology, 63*(MAY).

Dunker, K. (1929). Uber Induzierte Bewegung (pp. 180-259): Psychologische Forchung.

Eckhorn, R., Bauer, R., Jordan, W., Brosch, M., Kruse, W., Munk, M., Reitboeck, H. J. (1988). Coherent Oscillations: A Mechanism for Feature Linking in the Visual Cortex? (pp. 121–130): Biological Cybernetics.

Egly, R., Driver, J., & Rafal, R. D. (1994). Shifting Visual-Attention between Objects and Locations - Evidence from Normal and Parietal Lesion Subjects. *Journal of Experimental Psychology-General, 123*(2), 161-177.

Engel, A. K., Konig, P., Gray, C. M., & Singer, W. (1990). Stimulus-Dependent Neuronal Oscillations in Cat Visual-Cortex - Intercolumnar Interaction as Determined by Cross-Correlation Analysis. *European Journal of Neuroscience, 2*(7), 588-606.

Engel, S. A. (1994). fMRI of Human Visual-Cortex. *Nature, 370*(6485), 106-106.

Farrell, J. E. (1984). Visible Persistence of Moving Objects. *Journal of Experimental Psychology-Human Perception and Performance, 10*(4).

Farrell, J. E., Pavel, M., & Sperling, G. (1990). The Visible Persistence of Stimuli in Stroboscopic Motion. *Vision Research, 30*(6).

Feldman, J. (1999). The role of objects in perceptual grouping. *Acta Psychologica, 102*(2-3), 137-163.

Francis, G., Grossberg, S., & Mingolla, E. (1994). Cortical Dynamics of Feature Binding and Reset - Control of Visual Persistence. *Vision Research, 34*(8).

Frien, A., Eckhorn, R., Bauer, R., Woelbern, T., & Kehr, H. (1994). Stimulus-Specific Fast Oscillations at Zero Phase Between Visual Areas V1 and V2 of Awake Monkey. *Neuroreport, 5*(17), 2273-2277.

Gardner, J. L., Merriam, E. P., Movshon, J. A., & Heeger, D. J. (2008). Maps of Visual Space in Human Occipital Cortex are Retinotopic, not Spatiotopic. *Journal of Neuroscience, 28*(15), 3988-3999.

Greene, E. (1998). A Test of the Gravity Lens Theory. *Perception, 27*(10), 1221-1228.

Gruber, T., Muller, M. M., & Keil, A. (2002). Modulation of Induced Gamma Band Responses in a Perceptual Learning Task in the human EEG. *Journal of Cognitive Neuroscience, 14*(5), 732-744.

Haber, R. N., & Standing, L. G. (1970). Direct Estimates of Apparent Duration of A Flash. *Canadian Journal of Psychology, 24*(4).

Hammett, S. T. (1997). Motion Blur and Motion Sharpening in the Human Visual System. *Vision Research, 37*(18).

He, Z. J., & Ooi, T. L. (1999). Perceptual Organization of Apparent Motion in the Ternus Display. *Perception, 28*(7), 877-892.

Herzog, M. H., Otto, T. U., & Ogmen, H. (2012). The Fate of Visible Features of Invisible Elements. *Frontiers in psychology, 3*.

Humphreys, G. W., & Riddoch, J. (2007). How to Define an Object: Evidence from the Effects of Action on Perception and Attention. *Mind & Language, 22*(5), 534-547.

Irwin, D. E. (1991). Information Integration Accross Saccadic Eye-Movements. *Cognitive Psychology, 23*(3), 420-456.

Johansson, G. (1973). Visual Perception of Biological Motion and a Model for its Analysis. *Perception & Psychophysics, 14*(2).

Johansson, G. (1975). Visual Motion Perception. *Scientific American, 232*(6), 76-&.

Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The Reviewing of Object Files - Object-Specific Integration of Information. *Cognitive Psychology, 24*(2), 175-219.

Kalveram, K. T., & Ritter, M. (1979). Formation of Reference Systems in Visual Motion Perception. *Psychological Research-Psychologische Forschung, 40*(4).

Kasai, T., Moriya, H., & Hirano, S. (2011). Are Objects the Same as Groups? ERP Correlates of Spatial Attentional Guidance by Irrelevant Feature Similarity. *Brain Research, 1399*, 49-58.

Kawabe, T. (2008). Spatiotemporal Feature Attribution for the Perception of Visual Size. *Journal of Vision, 8*(8), 9.

Knapen, T., Rolfs, M., & Cavanagh, P. (2009). The Reference Frame of the Motion Aftereffect is Retinotopic. *Journal of Vision, 9*(5), 6.

Koffka, K. (1935). Principles of Gestalt Psychology. New York: Harcourt, Brace & World, Inc.

Kolers, P. A. *Aspects of motion perception*: Oxford: Pergamon Press.

Kramer, P., & Rudd, M. (1999). Visible Persistence and Form Correspondence in Ternus Apparent Motion. *Perception & Psychophysics, 61*(5), 952-962.

Kramer, P., & Yantis, S. (1997). Perceptual Grouping in Space and Time: Evidence from the Ternus Display. *Perception & Psychophysics, 59*(1), 87-99.

Lefton, L. A. (1973). Metacontrast - Review. *Perception & Psychophysics, 13*(1B).

Lin, Z. C., & He, S. (2012). Automatic frame-centered object representation and integration revealed by iconic memory, visual priming, and backward masking (vol 12, pg 1, 2012). *Journal of Vision, 12*(13), 1.

Livingstone, M. S. (1996). Oscillatory Firing and Interneuronal Correlations in Squirrel Monkey Striate Cortex. *Journal of Neurophysiology, 75*(6), 2467-2485.

Marino, A. C., & Scholl, B. J. (2005). The Role of Closure in Defining the ")bjects" of Object-Based Attention. *Perception & Psychophysics, 67*(7), 1140-1149.

Marr, D. (1982). Vision. . New York: W. H. Freeman.

McDougall, W. (1904). The Sensations Excited by a Single Momentary Stimulation of the Eye (Vol. 1, pp. 78-113): British Journal of Psychology.

McRae, K., Butler, B. E., & Popiel, S. J. (1987). Spatiotopic AND Retinotopic Components of Iconic Memory. *Psychological Research-Psychologische Forschung, 49*(4), 221-227.

Melcher, D., & Colby, C. L. (2008). Trans-Saccadic Perception. *Trends in Cognitive Sciences, 12*(12), 466-473.

Melcher, D., & Morrone, M. C. (2003). Spatiotopic Temporal Integration of Visual Motion across Saccadic Eye Movements. *Nature Neuroscience, 6*(8), 877-881.

Montangero, J., & Maurice-Neville, D. (1994). Piaget ou l'intelligence en marche. . Liège.: Mardaga.

Muller, M. M., Gruber, T., & Keil, A. (2000). Modulation of Induced Gamma Band Activity in the Human EEG by Attention and Visual Information Processing. *International Journal of Psychophysiology, 38*(3), 283-299.

Naito, S., & Cole, J. B. (1994). The Gravity Lens Illusion and its Mathematical Model. *Contributions to Mathematical Psychology, Psychometrics, and Methodology*, 39-50.

Nishida, S., Watanabe, J., Kuriki, I., & Tokimoto, T. (2007). Human Visual System Integrates Color Signals along a Motion Trajectory. *Current Biology, 17*(4), 366-372.

Ogmen, H. (1993). A Neural Theory of Retino-Cortical Dynamics. *Neural Networks, 6*(2).

Ogmen, H. (2007). A Theory of Moving Form Perception: Synergy Between Masking, Perceptual Grouping, and Motion Computation in Retinotopic and Non-Retinotopic Representations. *Advances in cognitive psychology / University of Finance and Management in Warsaw, 3*(1-2).

Ogmen, H., & Herzog, M. H. (2010). The Geometry of Visual Perception: Retinotopic and Nonretinotopic Representations in the Human Visual System. *Proceedings of the Ieee, 98*(3), 479-492.

Ogmen, H., Otto, T. U., & Herzog, M. H. (2006). Perceptual Grouping Induces Non-Retinotopic Feature Attribution in Human Vision. *Vision Research, 46*(19), 3234-3242.

Paakkonen, A. K., & Morgan, M. J. (1994). Effects of Motion on Blur Discrimination. *Journal of the Optical Society of America a-Optics Image Science and Vision, 11*(3).

Pantle, A., & Picciano, L. (1976). Multistable Movement Display - Evidence for 2 Separate Motion Systems in Human-Vision. *Science, 193*(4252), 500-502.

Patterson, R., Hart, P., & Nowak, D. (1991). The Cyclopean Ternus Display and the Perception of Element Versus Group Movement. *Vision Research, 31*(12), 2085-2092.

Petersik, J. T., Hicks, K. I., & Pantle, A. J. (1978). Apparent Movement of Successively Generated Subjective Figures. *Perception, 7*(4), 371-383.

Petersik, J. T., & Pantle, A. (1979). Factors Controlling the Complete Sensations Produced by a Bistable Stroboscopic Motion Display. *Vision Research, 19*(2), 143-154.

Petersik, J. T., & Rice, C. M. (2006). The Evolution of Explanations of a Perceptual Phenomenon: A Case History using the Ternus Effect. *Perception, 35*(6), 807-821.

Piaget, J., & Inhelder, B. (1969). *The Psychology of the Child*. New York: Basic Books.

Pikler, J. (1917). Sinnesphysiologische Untersuchungen.

Purushothaman, G., Ogmen, H., Chen, S., & Bedell, H. E. (1998). Motion Deblurring in a Neural Network Model of Retino-Cortical Dynamics. *Vision Research, 38*(12).

Ramachandran, V. S., Rao, V. M., & Vidyasag, T. R. (1974). Sharpness Constancy During Movement Perception - Short Note. *Perception, 3*(1).

Restle, F. (1979). Coding Theory of the Perception of Motion Configuration. *Psychological Review, 86*(1).

Ritter, A. D., & Breitmeyer, B. G. (1989). The Effects of Dichoptic and Binocular Viewing on Bistable Motion Percepts. *Vision Research, 29*(9), 1215-1219.

Rock, I., Auster, M., Schiffman, M., & Wheeler, D. (1980). Induced Movement Based on Subtraction of Motion From the Inducing Object. *Journal of Experimental Psychology-Human Perception and Performance, 6*(3), 391-403.

Rock, I., & Divita, J. (1987). A Case of Viewer-Centered Object Perception. *Cognitive Psychology, 19*(2).

Scholl, B. J., Pylyshyn, Z. W., & Feldman, J. (2001). What is a Visual Object? Evidence from Target Merging in Multiple Object Tracking. *Cognition, 80*(1-2), 159-177.

Sereno, M. I., Pitzalis, S., & Martinez, A. (2001). Mapping of Contralateral Space in Retinotopic Coordinates by a Parietal Cortical Area in Humans. *Science, 294*(5545), 1350-1354.

Shimozaki, S. S., Eckstein, M., & Thomas, J. P. (1999). The Maintenance of Apparent Luminance of an Object. *Journal of Experimental Psychology-Human Perception and Performance, 25*(5), 1433-1453.

Stewart, A. L., & Purcell, D. G. (1974). Visual Backward Masking by a Flash of Light - Study of U-Shaped Detection Functions. *Journal of Experimental Psychology, 103*(3).

Sun, J. S., & Irwin, D. E. (1987). Retinal Masking During Pursuit Eye-Movements - Implications for Spatiotopic Visual Persistence. *Journal of Experimental Psychology-Human Perception and Performance, 13*(1), 140-145.

Tarr, M. J. (1995). Rotating Objects to Recognize Them - A Case-Study on the Role of Viewpoint Dependency in the Recognition of 3-Dimensional Objects. *Psychonomic Bulletin & Review, 2*(1).

Tarr, M. J., & Pinker, S. (1989). Mental Rotation and Orientation-Dependence in Shape-Recognition. *Cognitive Psychology, 21*(2).

Tarr, M. J., & Pinker, S. (1990). When Does Human Object Recognition Use a Viewer-Centered Reference Frame. *Psychological Science, 1*(4).

Tassinari, G., Marzi, C. A., Lee, B. B., Di Lollo, V., & Campara, D. (1999). A Possible Selective Impairment of Magnocellular Function in Compression of the Anterior Visual Pathways. *Experimental Brain Research, 127*(4).

Ternus, J. (1926). Experimentelle Untersuchungen uber Phanomenale Identitat (Vol. 7, pp. 81-136): *Psychol. Forsch.*

Tootell, R. B. H., Hadjikhani, N. K., Vanduffel, W., Liu, A. K., Mendola, J. D., Sereno, M. I., Dale, A. M. (1998). Functional Analysis of Primary Visual Cortex (V1) in Humans. *Proceedings of the National Academy of Sciences of the United States of America, 95*(3), 811-817.

Tootell, R. B. H., Reppas, J. B., Kwong, K. K., Malach, R., Born, R. T., Brady, T. J., Rosen, B. R., Belliveau, J. W. (1995). Functional-Analysis of Human MT and Related Visual Cortical Areas using Magnetic-Resonance-Imaging. *Journal of Neuroscience, 15*(4), 3215-3230.

Tso, D. Y., & Gilbert, C. D. (1988). The Organization of Chromatic and Spatial Interactions in the Primate Striate Cortex. *Journal of Neuroscience, 8*(5), 1712-1727.

Tso, D. Y., Gilbert, C. D., & Wiesel, T. N. (1986). Relationship Between Horizontal Interactions and Functional Architecture in Cat Striate Cortex as Revealed by Cross-Correlation Analysis. *Journal of Neuroscience, 6*(4), 1160-1170.

von der Malsburg, C. (1981). The Correlation Theory of Brain Function. Gottingen: Max-Planck-Institute for BioPhysical Chemistry: Internal Report 81-2. [Reprinted (1994) in E. Domany, J. L. van Hemmen, & K. Schulten (Eds.), Models of Neural Networks II (pp. 95-119). Berlin: Springer-Verlag.

von der Malsburg, C. (1995). Binding in Models of Perception and Brain-Function. *Current Opinion in Neurobiology, 5*(4), 520-526.

Wallach, H. (1959). Perception of Motion. *Scientific American, 201*(1).

Wallach, H., Nitzberg, D., & Becklen, R. (1985). Vector Analysis and Process Combination in Motion Perception. *Journal of Experimental Psychology-Human Perception and Performance, 11*(1).

White, C. W. (1976). Visual Masking During Pursuit Eye-Movements. *Journal of Experimental Psychology-Human Perception and Performance, 2*(4), 469-478.

Yin, C., Shimojo, S., Moore, C., & Engel, S. A. (2002). Dynamic Shape Integration in Extrastriate Cortex. *Current Biology, 12*(16), 1379-1385.