Investigating the Effect of Texture Edges in Figure-Ground Segregation Using Psychophysical and Eye Tracking Experiments

Shumetha Kaur Sidhu, BSc.

Thesis submitted to the University of Nottingham

for the degree of Doctor of Philosophy in Psychology

December 2019

Abstract

Although it happens infrequently in the natural world, the human visual system is able to perceive objects defined solely by a difference in texture i.e. with no accompanying change in colour or luminance. However, studies on texture perception have frequently used figure-ground patterns with abrupt texture variations, with few studies using patterns with smooth texture variations. The work presented in this thesis considers the contribution of the edge and centre regions of a texture figure to orientation-based texture segregation. To this end, we used psychophysical (Chapters 4 – 6) and eye tracking techniques (Chapters 8 – 10) to investigate texture detection and segmentation. For the psychophysics experiments, the primary goal was to investigate if either an edge-based or region-based mechanism can account for **both** smooth and abrupt texture variations. We first examined this in Chapter 4, where we studied the effects of texture edges on figure-ground segregation. Lower thresholds were found when the texture figure had orientation contrast information at the edge **and** centre of the figure. Data modeling supports the notion that texture segregation involves a large-scale second-order texture filter i.e. akin to a region-based mechanism, but where information is extracted over a large albeit fixed-size region. Various studies were also conducted to determine what aspects of a texture figure would change the size of the second-stage filter. The size and aspect ratio of the figure were manipulated (Chapter 5), and also the spatial frequency of the texture pattern, age of the participants, and the viewing distance (Chapter 6). We found that higher spatial frequencies resulted in larger integration regions i.e. feeds into large second-stage filters, but age, viewing distance, figure size and aspect ratio did not influence the size of the integration region. For the eye tracking studies, the general aim was to investigate what information of a texture target is extracted in order to produce signals for eye movement control. We measured eye movements made by participants while

they searched for a texture figure embedded in a background. We found that irrespective of the types of orientation profiles, area-normalized data were that the centre region of a figure was looked at most often, and for longer durations. However, figures with information of orientation contrast at **both** the edge and centre of figure were easier to localise (Chapter 7), and produced the highest level of saliency in attracting eye movements (Chapter 10). In Chapter 9, we demonstrate that the visual system is also able to efficiently segregate a texture figure from the ground to accurately plan a saccade to the target figure, and these saccades are planned based on the representation of the whole figure shape as opposed to local salient regions. More specifically, saccades were directed to the centre of gravity of the target, with some degree of undershoot. Finally, the similar size of the integration region for the eye tracking (Chapter 10) and psychophysics experiments implies that the saccadic system receives input from the mechanism that segregates figure-ground texture stimuli.

Acknowledgements

I would like to express my sincere gratitude to my supervisor, Dr David Keeble, for not only taking a chance on me, but also patiently mentoring and guiding me throughout my entire PhD. Despite your extremely hectic schedule, you always made time for me, and also went out of your way to help out with other non-research related matters. It is also thanks to your support and encouragement that I was able to experience many professional milestones, and I truly could not imagine having a better supervisor than you. I am also grateful to my other supervisor, Dr Harriet Allen, for the opportunity she gave me to conduct research with older people. Not only was it extremely enjoyable, but it also opened my eyes to real-world applications of our line of research. To both David and Harriet, I am so very grateful for everything you have done for me that has lead me to this point in my life.

I would like to also thank Dr Jessica Price for being a ray of sunshine in our jungle. Your assurance that "it will all be fine" helped tremendously throughout my undergraduate degree, and especially so during my graduate degree. A very big thank you to Dr Alex Estudillo and Dr Jason Bell for being my examiners. Alex, I am really glad that you have been with me from the first progression review right up to the VIVA. Many thanks also go to my fellow graduate students in BB47, past and present, for the company along this interesting journey of ours. It was great having people to commiserate with.

Special thanks to Dr Jonathan Pierce for his very helpful suggestions and samples of codes to make our experiment more efficient. To all my observers that spent hours and hours on my psychophysics studies, I am very thankful for your participation and perseverance. To all the other participants, I am also very appreciative of all your efforts. To my elderly participants in particular, thank you for the very enjoyable experience! To Nabihah, Anis, and Zahira – thank you for your assistance in data collection.

Finally, I would like to thank to my family. In some manner or another, you have all helped shaped me into the person that I am now. Undertaking this PhD without your support would have been extremely difficult – so thank you for making this easier on me. To my parents, especially my mum, thank you for giving me the best opportunities that you could, and most importantly, for the unconditional love over the years. To my sister, thanks for always having my back.

Table of Contents

List of Figures

List of Tables

Chapter 1

General Introduction

Texture refers to the attributes of the composing elements that represent the characteristics of an object. These composing elements have to be comparatively small compared to the size of the object so that these elements represent the texture of the object, as opposed to the shape of the object i.e. the "stuff" in the image (Landy & Graham, 2004). Although we are mostly unaware of it, our visual system relies on texture information to identify objects, distinguish objects from the background, unify surfaces, and provide information about depth (Bergen, 1991). Our present interest here is not the visual perception of surface texture (i.e. smooth, rough, glossy), but the perception of visual texture as defined by an isolated feature (e.g. motion, orientation, contrast).

Moreover, a phenomenon of particular interest in visual texture perception is texture-based segregation (see Figure 1.1 for an illustration). As the name implies, it is the segregation of regions of texture based on feature characteristics present in the texture. The ×-shaped elements on the left side of the texture stimulus appear as a distinct region rapidly with no conscious effort or search. However, the T-shaped elements on the right side do not appear distinct from the background. While it is true that texture segregation based purely on an isolated feature is not a very common occurrence in everyday life, the contribution of texture properties in the processing of complex scenes cannot be neglected.

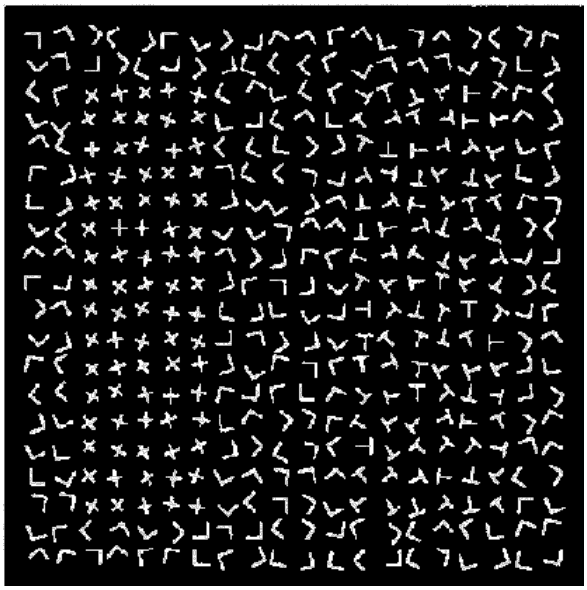


Figure 1.1. Example of texture-based segregation. The region on the left comprises X-shaped elements embedded in L-shaped background elements, which segregates easily. The region on the right, which comprises T-shaped elements embedded in L-shaped background, does not appear as a distinct region. From Bergen (1991).

While there are many features of textures that can be studied (e.g. colour, luminance, motion, contrast), the work presented in this thesis focuses on the feature of orientation. The interest in orientation-based texture segregation is because orientation plays a role in perceptual organisation (Julesz, 1981), and it also has a strong neurophysiological basis (Hubel & Wiesel, 1977). Although the visual world is very rarely defined by a sole texture property, studying an isolated property of texture, in our case orientation, will allow us to gauge its role in processing complex visual scenes.

The experiments presented in this thesis investigate orientation-based texture segregation. Specifically, we are interested in the role texture edges (a region where a texture property changes abruptly) play in the segregation process. This is because of the discrepancies in the literature, where some studies suggest that texture edges are crucial for figure-ground segregation (e.g. Landy & Bergen, 1991; Wolfson & Landy, 1998), while others have shown that texture segregation is still possible in the absence of a texture edge (e.g. Kingdom &

Keeble, 1996; Ben-Shahar & Zucker, 2004). Additionally, it will become clear as we review the literature further on that an abundance of studies that investigate texture perception typically use stimuli with abrupt (or sharp) texture edges, without considering the effects of stimuli with smooth texture variations.

To investigate orientation-based texture segregation, we have used two different types of experimental methods: psychophysics and eye tracking. A key point to note about this is that the overarching objective for all the experiments (both psychophysics and eye tracking) is to investigate whether or not texture edges are crucial for figure-ground segregation. We used two different methodologies to investigate this as we believe that the psychophysics studies, while extremely important and informative, limit the movement of participants' eyes. We hope to **add** to our understanding of the existing literature by also using eye tracking measures to investigate the role of texture edges in figure-ground segregation.

This thesis will have two parts to it, first for the psychophysical studies, and subsequently the eye tracking studies. The first introductory chapter (Chapter 2) is what builds up our rationale of investigating the role of texture edges in figure-ground segregation, and also justifies the types of stimuli we created. The second introductory chapter (Chapter 7) on the other hand is more specific to the eye tracking studies.

Chapter 2

Introduction to Texture Perception

2.1 Early Work in Texture Segmentation and Psychophysical Experiments

Some of the earliest influential work in the field of texture segregation comes from Bela Julesz and Jacob Beck. Julesz (1962) used patterns composed of two fields, with each field being described by a probability distribution. The first order probability distribution is the probability of a dot element having a certain luminance. For example, one field would have a 5/8 probability of being black, and 3/8 probability of being white, and vice versa for the other field. This reliably produced segmentation. Second order probability distribution is the probability that the endpoints of a dipole, which could be of any length and orientation, would fall on certain luminance combinations of the stimuli. Therefore, first order statistics take into account only individual pixels, while second order statistics account for the interaction of any pairs of pixel elements. As a result, pattern segmentation can occur when there is identical first order statistical distribution, but different second order distributions. However, if both first and second order distributions are identical, segmentation will not occur. Third order probability distribution is the probability that all 3 vertices i.e. the 3 corners of a 3-sided polygon, would fall on the same luminance. This was achieved by having one half of the field having a higher probability of the 3 vertices falling on the same colour, while the other field had a lower probability of this occurring. When patterns were formed using fields that had the same first and second order distribution, but different third order distributions, segregation was very weak (see Figure 2.1).

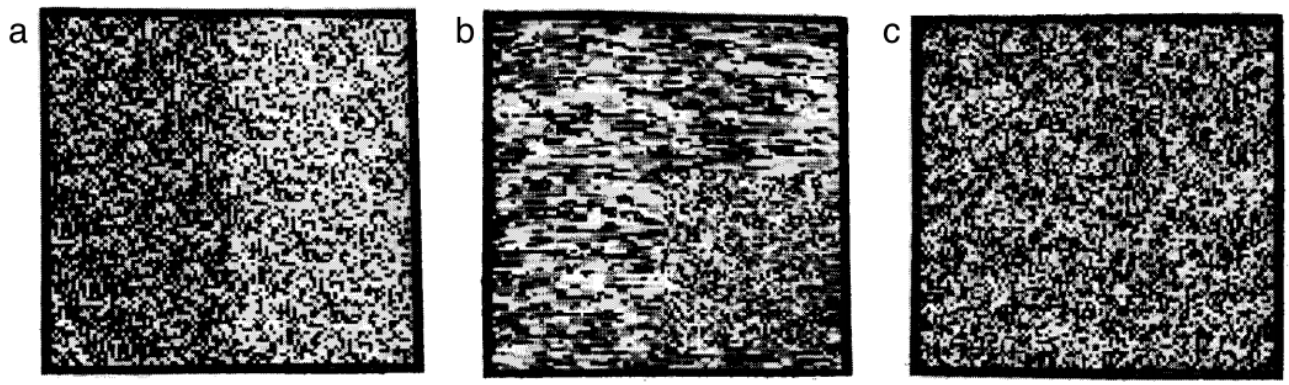


Figure 2.1. a) Two fields with different first order probability distributions, b) Two fields with identical first order probability distribution, but different second order probability distribution, c) Two fields with identical first and second order probability distribution, but different third order probability distribution. From Julesz (1962).

Caelli and Julesz (1978) subsequently proposed that feature detection was carried out by two different visual pathways - which they term Class A and B. Class A was important for texture segregation as they consisted of dipole detectors tuned to specific lengths and orientations. The proposed that texture segregation occurs when either pathway detects a difference in input between multiple regions. It was noted that segregation of texture could not be fully accounted for by these dipole detectors as they were only able to segregate a subset of stimuli.

Julesz (1981) claimed that visual perception is achieved by an attentive and pre-attentive stage. In the pre-attentive stage, despite the lack of complex form processing, there is an effortless extraction of local image features. These features, called textons, are used to pre-attentively discriminate textures. It was found that even when an image had different second order statistical distributions due to an embedded pattern (Figure 2.2b), the image appeared continuous, and did not segregate. However, when the embedded pattern contained striped textons (Figure 2.2d), the image was segregated. They suggest that segregation occurs not solely on second order distribution, but is strongly reliant on the density changes in textons.

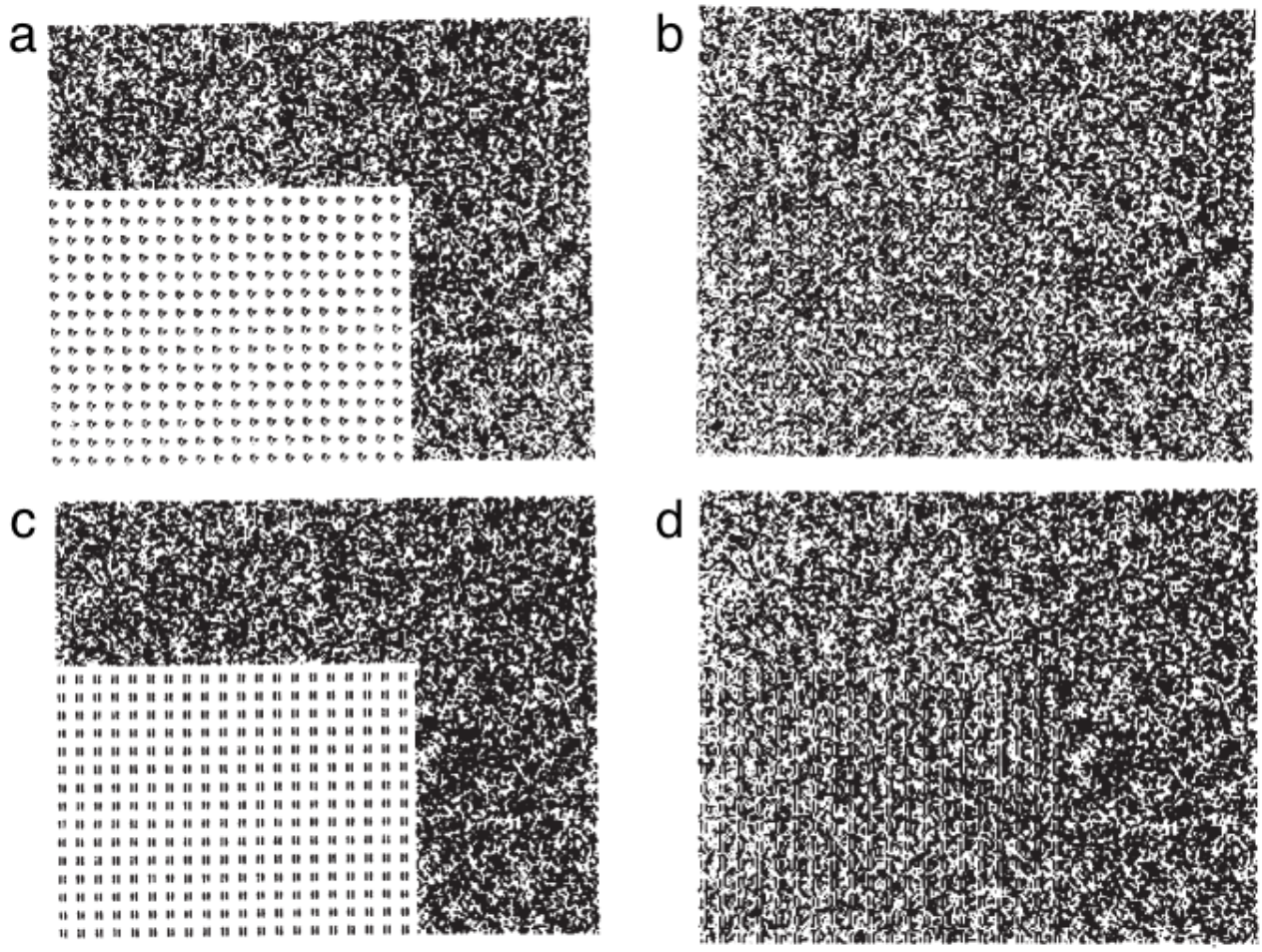


Figure 2.2. a) Pattern that contains no textons is embedded into a randomly dotted area, b) Same as A, but with random dots filling the blank gaps, c) Pattern that has textons is embedded into a randomly dotted area, d) Same as C, but with random dots filling the blank gaps. Fig. 2a and 2c are included for illustrative purposes, and were not used as experimental stimuli. From Julesz (1981).

Beck's work on the other hand had a different approach to that of Julesz. While Julesz was more focused on what caused segregation, Beck (1966) focused on what caused elements to group together. Elements that are similar to each other would be grouped together, and be perceived as a distinct entity from elements that fail to group together. Participants were required to indicate where in a 3-part pattern stimulus the natural break occurs (where the elements failed to group together), and they often indicated a natural break in the pattern where the orientation change occurred (i.e. upright T vs. slanted T). Only in the absence of

orientation change did line arrangements influence break in grouping. To further this Beck (1982) stated that simple properties (e.g. line, movement, size, colour, brightness, etc.) present in an image aid in the segmentation process. He proposed a model that; 1) detects features by receptive fields, 2) features extracted are linked in terms of the feature elements, 3) features of textural elements that are nearby are compared, and differences are encoded, 4) decision units take into account the differences encoded, and as a result segment the display. In later work, Beck, Sutter, and Ivry (1987) found that segregation is not about the elements or features within a texture, but the difference in stimulation experienced by spatial frequency channels based on different texture regions.

Contrary to Beck's findings of feature linking, Ziebel and Nothdurft (1995) found that feature linking does not reliably produce a pop-out region of segregation. This was in line with Nothdurft's (1985a,b) early work in texture segregation, in which it was found that texture boundaries are more important than grouping of similar elements. Nothdurft (1985a) found that line length and line spacing played a role in performance of a discrimination task (global patterns with number 0-9). When line length was held constant, the performance of participants improved when the line spacing decreased. Likewise, when line spacing was held constant, performance improved when the line length was increased. Nothdurft (1985b) also found that line length improved performance in a discrimination task (horizontal and vertical bars), as well as an orientation sensitivity task (make estimates of the line orientations according to 1 of 8 classes). Participants' performance of the orientation sensitivity task was better compared to the discrimination task.

Nothdurft (1985a) found that texture segregation was not only affected by the structure density of a stimulus (number of elements per unit area), but also the orientation contrast. At low orientation contrast (20°), increasing the line length improved performance. However,

increasing line length did not improve performance when orientation difference was very low (10°). With higher orientation contrast (30° and 90°), only a minimal increase in line length was needed to improve performance, before a plateau was reached. Nothdurft took the dependency of segregation performance on orientation contrast as well as the structural property of the stimulus to be due to the computation of gradients. The structural gradient takes into account the spacing of the elements, and the ability to segment a stimulus with a certain orientation drops when elements are placed further apart. Nothdurft (1991) also found that increasing orientation variation of the line elements resulted in poorer performance. That is, increased variation to the line elements resulted in higher threshold values of orientation contrast of figure and background.

2.2 Neurophysiology of Texture Perception

Texture segregation and the related processes associated with it have been extensively studied not only through psychophysical methods, but also physiological methods. While the experiments presented in this thesis do not employ any of these physiological methods, they are nonetheless interesting as they provide additional insights into the neurological correlates that process texture segregation. As such, the following section will focus on texture perception studies that use single-cell recording of monkeys, EEG, and fMRI techniques. Furthermore, of these neurophysiological methods, the single-cell recordings and EEG studies are particularly informative with regards to the time course of activation for the visual processing of texture targets.

Our interest in this time course of activation lies primarily in the fact that we are interested in investigating the contribution of the edge/boundary of a figure, and also the centre/surface of the figure. As we will discuss later in Section 2.6 (Filling-in Hypothesis), scene segmentation does not occur instantaneously throughout the entire figure, but actually gets “filled in” in a

gradual manner. However, the studies that will be discussed in Section 2.6, while informative about the effects of filling-in, use the masking paradigm in their experiments. The studies discussed in this section on the other hand do not use such a method, and also have the added benefit of tracking the time course of neuronal activity in response to specific regions of a texture stimulus.

Therefore, when we use stimuli with different information of orientation contrast at the edge and centre region, we will be able to compare performance to the various different stimuli types and stimulus duration in our experiments to the time course of neuronal activity as discussed here. This will be discussed further in Chapter 4, where we present an experiment to specifically investigate if edge extraction precedes surface filling-in.

The reader is also directed to Chapter 7, where we discuss in greater detail the human visual system and neural control of eye movements. Furthermore, Section 7.5.3 also discusses the strength of neuronal responses associated with saccades made to figure-ground textures, which is closely related to what we will discuss in Section 2.2.1 below.

2.2.1 Single-cell Recordings

Lamme (1995) recorded neuronal activity in macaque monkeys while they viewed figure-ground texture stimuli defined by orientation or direction of motion contrast. The local stimulation to the receptive field (RF) was always the same, and the only difference was whether the ‘figure’ or the ‘ground’ region of the stimuli would fall on the RF. Results showed that there were increased neuronal responses when the RF was on the figure compared to the ground, and this was the case for elements beyond the receptive field (RF), which suggest the use of contextual information i.e. information outside the classical receptive field. Interestingly, neuronal responses diminished as soon as the RF was on the

background, even if the RF was close to the border. When the RF was on the figure, enhanced responses occurred approximately 30-40ms after neuronal responses to stimulus onset.

Subsequently, Lamme, Rodriguez-Rodriguez, and Spekreijse (1999) explored the temporal aspects of figure-ground segmentation in macaque monkeys. They used texture stimuli that had 90° orientation contrast, with the size of the square figure spanning 4° of visual angle. They found processing of texture stimuli begun with local feature detection, which led to the detection of feature boundaries, and finally a representation of figure and background surfaces. They were interested in neuronal activity in V1 when the RFs were inside and outside of the figure. As before, they found that neuronal responses to RFs on the figure were enhanced compared to the background. To gauge the temporal aspects, they looked at 10ms slices of responses. After stimulus onset, neuronal responses to the stimulus began at 20-30ms, and figure-ground responses only began at 70-80ms (with strongest responses at RF near boundaries). At 90-100ms, boundary responses continue to increase while no changes have occurred for figure centre. At 115-125ms, peak responses to boundaries are observed, with only a slight response to figure centre. Responses for the whole figure surface (boundary + figure centre) only appear at 150-160ms (see Figure 2.3). These findings show that boundary/edge detection precedes and initiates the process of figure surface filling (also see Section 2.6: Filling-in Hypothesis).

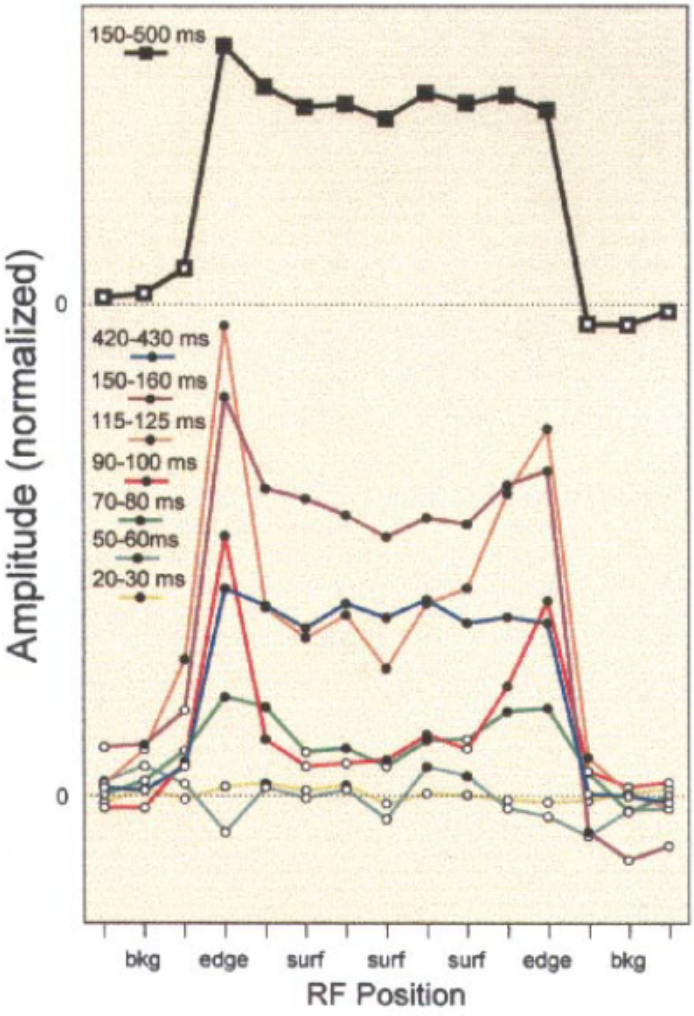


Figure 2.3. (Top) The difference of amplitude between figure and background responses between 150 and 500ms. (Bottom) The difference of amplitude between figure and background responses in 10ms slices. From Lamme, Rodriguez-Rodriguez, and Spekreijse (1999).

Contrary to the findings of Lamme and colleagues, Rossi, Desimone, and Ungerleider (2001) suggested that neurons in V1 do not segregate figures from backgrounds, but they do signal changes in texture boundaries. Rossi and colleagues used single-cell recordings of Rhesus Macaques viewing texture stimuli of black lines on a grey background that had an orientation contrast of 90°. By manipulating the size of the circular figure patch, they could record the RF neuronal responses to figure elements, while increasing the distance of the boundary from the RF. They found that figure responses were only enhanced when the figure size was 1°,

but not 2° or 4°. This was because at 1° figure size, the boundary was close enough to the RF for contextual effects to aid responses. This also held true for V2 responses, where enhanced figure response was found when the figure size was 1° and 2° only. In addition, they manipulated the stimulus configuration and included stimuli that contained an oriented textured figure and grey backgrounds, and grey figure and oriented textured backgrounds and repeated the study. As before, they found enhanced responses to figures which were 1° and 2° in size. Of particular interest is the stimuli that had a blank grey foreground and oriented textured background - in this case, the grey figure with no orientated elements should theoretically have decreased responses, but they still found an enhanced response. This gives support to the idea that figural elements are not causing the enhanced neuronal responses in V1, but the boundary between figure and background due to discontinuities in texture properties are resulting in enhanced response in V1.

Poort et al., (2012) used single cell recordings of area V1 and V4, and found that boundary/edge detection preceded region filling-in, which is in accordance other previous research (e.g. Lamme et al., 1999; Caputo & Casco, 1999). Like the previous studies, Poort and colleagues used stimuli of textures that had a patch in the centre with an orientation contrast of 90°, and the monkeys had to make a saccade to the target position. However, their study focused on how attention modulated figure enhancement (figure-ground modulation, FGM). When looking at FGM when the RF was on the figure centre, figure edge, and background, they found that there was enhanced FGM to figure centre and edge compared to the background. The increased FGM responses to the edge peaked at an earlier time compared to the centre response. When looking at attentional effects, when the stimulus was attended, there was enhanced response to the edge, followed by enhanced response to the figure centre. When the stimulus was not attended, there was an enhanced response to the

edge, followed by diminished response to the figure centre - which suggests only a partial filling-in of the figure centre in the unattended task.

Poort and colleagues also looked at V4 responses. They found that with both attended and unattended tasks, there were enhanced responses to the figures. However, while the pattern of activation was the same initially, at later times, the response to the figure was stronger in the attended task compared to the unattended task. Also to note, peak responses to the figure were at 175ms in V4 neurons, which is approximately 50ms before peak response to the figure in V1 neurons. A timeline of neuronal activation is as follow; 1) visual response to stimulus onset in V1 and V4 at approximately the same time, 2) V1 edge modulation at 60ms, 3) V4 figure-ground modulation at 67ms, 4) V1 centre modulation at 95ms. All this taken together suggests that edge detection is an early and automatic process (as seen by responses to edges in both attended and unattended tasks), while region filling-in depends on feedback from higher V4 region, which is modulated by attention. Another interesting find is that increased FGM responses by neurons in V1 and V4 could accurately predict the landing position of the saccade (i.e. if FGM was stronger to the right or left, the saccade made to the figure was respectively made to the right or left side of the figure), which suggests that V1 and V4 also play a role in saccade planning. Poort and colleagues put forth a model that is discussed in Section 2.7.

2.2.2 EEG Studies.

Our exploration into EEG studies of texture segmentation will focus on the temporal aspects of texture segmentation, the effect collinearity and attention have on texture segmentation, and breaking down texture segregation into boundary detection and surface segregation.

Using an fMRI based EEG source-imaging technique, it was found that an orientation defined texture stimuli that was segmented had increased activity in V1 compared to uniformly oriented stimuli. This increase in V1 activity begins at around 143ms (Appelbaum, Ales, & Norcia, 2012). Lachapelle, McKerral, Jauffret, and Bach (2008) compared the processing speed of low-level features (such as orientation) and texture segregation using Visual Evoked Potentials (VEPs). They found that after stimulus onset, texture segregation VEPs (tsVEP) were not as fast as low-level feature VEPs (llVEP). This is in accordance with Nothdurft (2000) who found that compared to the time is takes to detect a single line, longer response times were needed to detect a pop-out region. It implies that the segregation of a texture region is a time consuming process that operates on a slower mechanism than a one that extracts low-level features.

It has also been shown that collinearity increases the detectability of a pop-out region (e.g. Harrison & Keeble, 2008; Li, Tseng, & Li, 2013). Caputo and Casco (1999) found evidence of a collinearity edge effect, where elements that are parallel to the border edge are more salient compared to elements that are orthogonal. In their study, they found tsVEP had peak responses at approximately 20ms faster for collinear edges.

Subsequently, Casco, Grieco, Campana, Corvino, and Caputo (2005) studied the effect attention had on this collinearity edge (configural) effect. They manipulated the configuration of the stimuli, while holding orientation change constant (see Figure 2.4b and c). Behavioural and VEP data were collected under 3 attentional conditions - explicit attention, implicit attention, and unattended. For all 3 attentional groups, the participants had to judge the orientation (left or right tilted) of the bar. In one group, participants had to judge the orientation of the bar for every trial (explicit attention). In the other two groups, the participants only had to make a judgement of orientation on the last stimulus presented to

them. However, one group was passively viewing the stimulus (implicit attention), while the other was engaged in a secondary number discrimination task (unattended). With explicit attention, parallel configuration had better performance compared to orthogonal configuration. This was also demonstrated in tsVEP, whereby parallel configuration had increased V1 activity in early N75, P100, N150, and N200. With implicit attention, no configural effect was observed in the behavioural and VEP data, but texture segregation based on orientation difference was still possible. In the unattended condition, no configural effects were present, and segregation did not occur. Taken together, this implies that attention facilitates the texture segregation process not by orientation changes, but by grouping operations used.

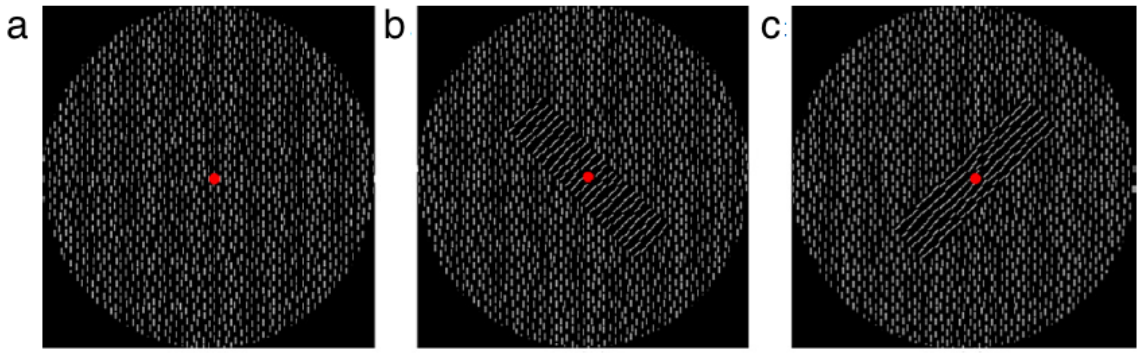


Figure 2.4. a) A uniform texture. A texture bar is segregated from the background by an orientation difference of 45°. b) Parallel configuration, where the short-edge boundary has the same orientation as its line elements. c) Orthogonal configuration, where the short-edge boundary has different orientation to its line elements. From Casco et al., (2005).

Finally, Scholte, Jolij, Fahrenfort, and Lamme (2008) found that boundary detection preceded surface segregation. In their study, they pooled EEG electrodes that were spatially close to produce five regions of interest: occipital, peri-occipital, temporal, parietal, and frontal regions. Using 3 different types of stimuli (see Figure 2.5), they were able to differentiate the neuronal activation for surface segregation and boundary detection + surface segregation. Differing responses between the stack and frame stimuli are attributed to surface segregation

mechanisms, while differing responses between stack and homogenous textures, and frame and homogenous textures are attributed to both surface segregation and boundary detection mechanisms. It was found that texture boundary detection occurs at the early visual cortex (occipital area) and spreads in a feedforward manner. Responses were first observed at the occipital area at 92ms, with activation spreading to the peri-occipital area (104ms), then to the temporal area (104-108ms) and then parietal area (104-120ms). For surface segregation, activity first began at the temporal area at 112ms, and spreads to the peri-occipital area (140ms), and then to the parietal and occipital areas (172ms). Activation due to surface segregation and boundary detection appears in the frontal areas after 208ms. It is important to note that in all areas except the frontal area, surface segregation always occurs at a later time. Caputo and Casco (1999) found activation in occipital areas at 140-160ms for boundary detection, and 200-260ms for surface segregation, which is later compared to that found by Scholte and colleagues, which were 92ms and 172ms respectively.

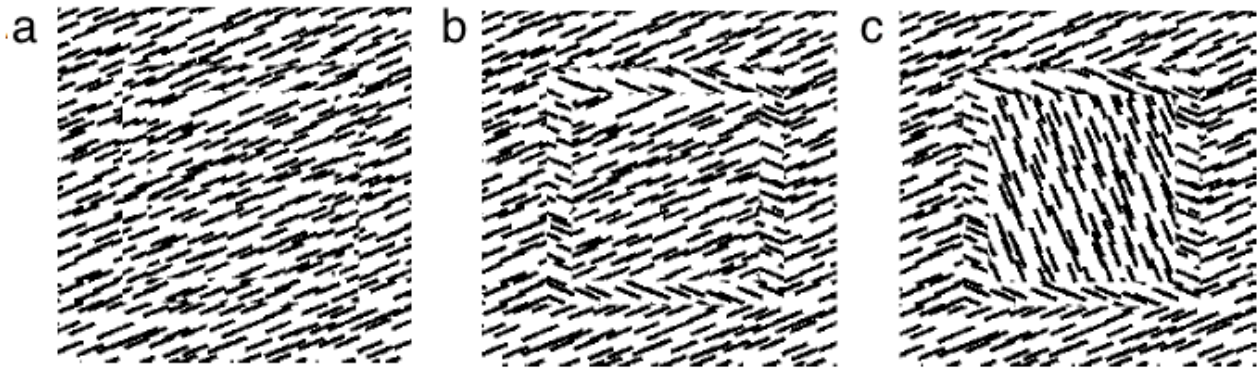


Figure 2.5. The 3 types of stimuli used in the experiment by Scholte and colleagues. a) Homogenous textures with uniformly oriented lines, b) Frame texture, with a 'frame' of different orientation imposed onto the background, c) Stack texture, with a 'frame' imposed onto the background. The orientation within the frame is also different from the background and frame for the stack texture. Differing responses between stack and frame stimuli are attributed to surface segregation mechanisms, while differing responses between stack and homogenous textures, and frame and homogenous textures are attributed to both surface segregation and boundary detection mechanisms. From Scholte et al., (2008).

2.2.3 fMRI Studies.

The temporal aspects in fMRI studies are not quite good enough to measure the precise timing of activations like EEG signals, but through blood oxygen level dependent (BOLD) activity, it is possible to track signal strength across the visual hierarchy.

Contrary to findings of V1 involvement in texture segregation (e.g. Lamme at al., 1999), Kastner, De Weerd, and Ungerleider (2000) did not find any activation in V1 or V2 in response to texture segregation. Instead, they found that a checkerboard pattern of orientation discontinuities evoked enhanced BOLD responses in V4, TEO, and slight activation in V3A. They postulated that the higher areas played an important role in texture segregation as these areas have large receptive field that can guide the segregation process. Despite the fact that previous studies have found V1 responses to figure-ground stimuli (e.g. Lamme, 1995), Kastner and colleagues believe that higher visual areas like V4 mediate responses in V1 and V2 via feedback connections, and this is demonstrated by the fact that neurons in V1 responded to stimuli outside of their small receptive field.

The results of Kastner et al., (2000) which show no activity in V1 during texture segregation is puzzling considering the other findings in the literature (e.g. Lamme et al., 1999' Rossi et al., 2001; Casco et al., 2005). In response to this, Scholte et al., (2008) claimed that the reason Kastner and colleagues failed to find V1 and V2 activation during texture segregation was the sample size used and attentional factors. To note, Kastner's study did not require participants to attend to the texture stimuli. That is, participants (N=5) were required to maintain fixation at the centre of a screen, and perform a counting task, while the texture stimuli were displayed to the upper right hand side of the fixation stimulus.

In contrast to Kastner et al., (2000), Scholte et al., (2008) explicitly told participants (N=25) in advance that there will be different stimulus configurations (see Figure 2.5) presented while they were performing a task that required them to foveate on the fixation point i.e. participants did not look at the texture figures, but were aware of their presence. They found that boundary detection and surface segregation occur in all early visual areas, including V1, and that early processing is not limited to boundary detection. Areas V1, V2, V3, V4, and V3a were involved with both boundary detection and surface segregation, while areas PPA and FFA responded to scene segmentation, but not boundary detection. According to BOLD activation, boundary signals get progressively stronger as they pass through the visual hierarchy, while surface segregation activation remains constant throughout all visual areas. They noted a pattern of feedforward detection of boundaries, and feedback connection for surface segregation.

Thielscher, Kolle, Neumann, Spitzer, and Gron (2008) also found strong BOLD activation in area V4 during texture segregation. In their study that looked at behavioural responses as well as pre-attentive BOLD activation, they found that in accordance with the literature, increased orientation contrast resulted in enhanced performance. Based on BOLD activation, they found that increased orientation contrast significantly increased BOLD activation in V4, VP and LOC (mid-level areas), but not V1d, V1v, V2d, V2v, and V3 (early visual areas). Thielscher and colleagues suggested that texture segregation is guided by pre-attentive signals from the mid-level visual areas. They propose a model of texture segregation, which is described in Section 2.7.

Based on the literature discussed thus far, there is plenty of support for the involvement of V1 in texture segregation. There are also some studies (e.g. Rossi, Desimone, & Ungerleider, 2001; Poort et al., 2012; Scholte et al., 2008; Thielscher et al., 2008) that suggest the

involvement of V2 and V4, with most of these suggesting that feedback from the higher areas are involved via a feedback connection to enhance texture segregation. However, the involvement of these regions in the process of texture segregation is not certain. Some studies suggest that the involvement of V1 cells is to identify the figure from the backgrounds, while others suggest that it signals boundaries between the figure and background.

Most importantly however, is the rather consistent finding that boundary detection precedes surface segregation. Despite that, the time course of neuronal activation in response to boundary detection and surface segregation does differ between studies e.g. 140-160ms and 200-260ms (Caputo & Casco, 1999), 92ms and 112ms (Scholte et al., 2008), 60ms and 90ms (Poort et al., 2012). In Chapter 4, we present an experiment, in which one of the objectives was to investigate if texture segregation is mediated by an edge extraction process that precedes surface segregation, and how, if at all, this is affected by the stimulus duration.

2.3 Configural Effects on Texture Segregation

Even though the work presented in this thesis does not specifically investigate how the configuration of elements can mediate texture segregation, it is important to note that the configuration of elements is in fact able to influence texture segregation. Thus, in the interest of investigating segregation of regions based **purely** on the texture properties of an image, consideration has to first be given about what other effects can influence segregation performance. In this section, we will discuss several studies that investigate how the configuration of elements can mediate texture segregation. In light of these findings, we will highlight at the end of this section a few ways in which we will perform control checks in our experiments to ensure that any outcome we have is not due to any of these configural effects.

Field, Hayes, and Hess (1993) identified several conditions that facilitated the identification of a path of Gabor patches from a random background array of Gabor patches. They found that a path was easily segregated from the background if; 1) the path is placed end-to-end (collinear) instead of side-to-side (Figure 2.6a vs. 2.6c) 2) the elements that constituted the path had smaller orientation difference from each other (Figure 2.6a vs. 2.6b), and 3) the elements were spaced closer together. They proposed the use of association fields to aid visual processing. The association field, which is wider than the classical receptive fields, is the region around an element that groups together elements based on association strengths.

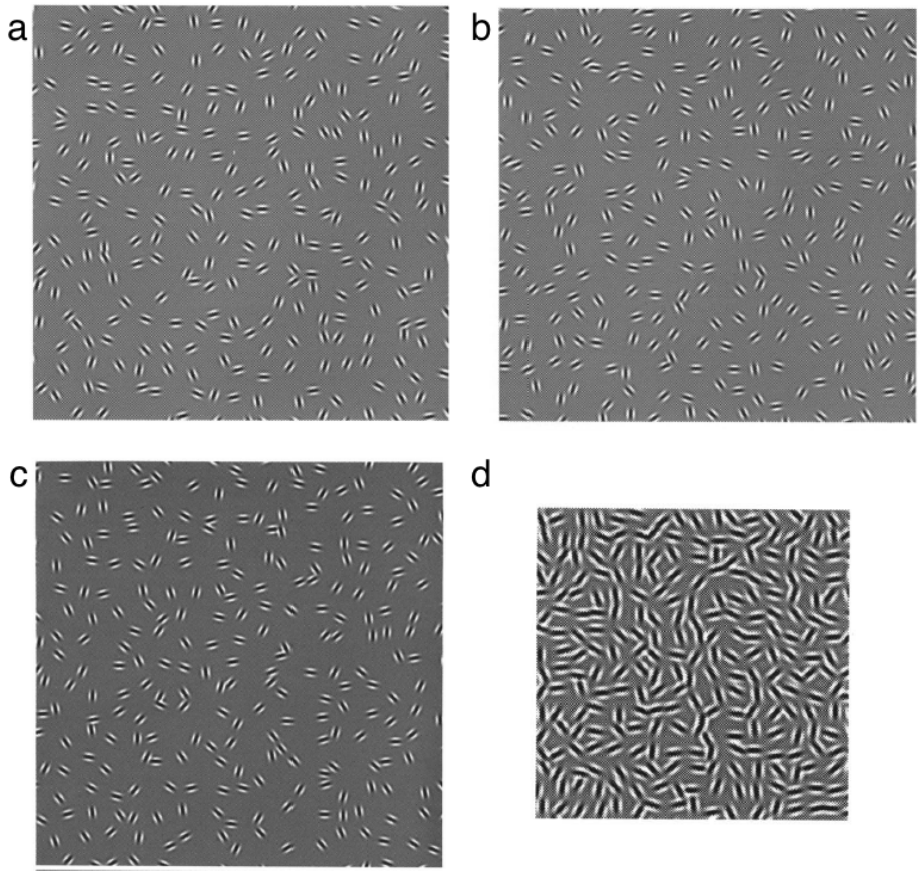


Figure 2.6. a) End-to-end placement of path, with each placement of element along the path having an orientation difference of 30°. b) Same as A, but with an orientation difference of 60° between elements of the path. c) Side-to-side placement of path, with each placement of element along the path having an orientation difference of 30°. d) Stimuli with 0.25 degrees spacing. From Field, Hayes, and Hess (1993).

Additionally, Wolfson and Landy (1995) found that texture segregation was not only influenced by the texture gradient, but also by the configuration of elements along the texture defined edge. They had participants perform a two-interval forced choice task to discriminate which interval had a straight as opposed to a wavy edge. They found that participants had enhanced performance when the elements on either side of the edge were parallel to the orientation of the edge boundary rather than orthogonal to the orientation of the edge boundary. Li and Li (2008) also found an increased saliency for elements along the border that were parallel in orientation to the orientation border. They had participants perform a change detection task, and found that participants were better at detecting the change in the parallel configuration (elements along figure border were parallel to the border) compared to the orthogonal configuration (elements along figure border were orthogonal to the border). According to the V1 model (e.g. Li, 1999a), which will be discussed further in Section 2.7, the parallel border bars had increased saliency due to the collinear facilitation from the neighbouring collinear elements. The orthogonal borders that have no collinear neighbours have no facilitation, hence the decreased saliency and poorer performance.

In addition, Giora and Casco (2007) found that performance in a binary classification task improved when edge elements were parallel to the edge and collinear (arranged end-to-end in a snake-like pattern) compared to when they were non-parallel and non-collinear. They also investigated how the arrangement of elements (either parallel or non-parallel to the edge) in the texture region affected performance. They found that temporal thresholds were lower by 22.5ms for the non-parallel condition compared to the parallel condition. A subsequent experiment found improved performance for the non-parallel condition even when the orientation contrast was reduced from 90° to 45°. Therefore, at the texture edge, facilitation occurs when elements are collinear and parallel to the edge, but at the texture region, facilitation occurs when elements are collinear and non-parallel to the edge.

Robol, Grassi, and Casco (2013) had participants perform a discrimination task, whereby participants had to indicate if the texture-edge was horizontal or vertical. They found that performance was enhanced when the region near the edge consisted of elements that were parallel (iso-orientation) compared to non-parallel (ortho-oriented) to the edge orientation. The reverse was true for elements away from the edge, in which performance was improved when elements away from the edge were non-parallel compared to parallel to the edge. They termed this respectively the iso-near and ortho-far effect. They also found evidence that suggests that the ortho-far contextual effects involve 1st order filters. They did this by including variations that are sensitive to first-order filters (the neural correlates of these filters are said to be neurons in V1, which are sensitive to orientation, contrast polarity, and positioning) to the far-region of the texture, and found that the ortho-far facilitation disappeared.

It has also been found that collinearity (Harrison & Keeble, 2008) and co-circularity (Motoyoshi & Kingdom, 2010) play a role in texture discrimination. Motoyoshi and Kingdom (2010) found that compared to random texture arrangements, textures with increased co-circularity (elements tend to form a circular arrangement) are discriminated with greater ease. This is the case for a wide range of orientation contrast, suggesting an important role co-circularity plays in texture segregation. As for Harrison and Keeble (2008), they found that by increasing the collinearity of texture elements, performance in a discrimination task improved.

Jingling and Tseng (2013) conducted a study in which a tilted bar was superimposed on an element within an array (red circles in Figure 2.7a and b), and participants were tasked with detecting the location of the superimposed bar. They found that when the bar was superimposed on a salient (column with a different orientation to the rest) collinear column,

performance was slower compared to when the bar was superimposed on a collinear non-salient element. Li, Tseng, and Li (2013) subsequently found that collinear columns were not highly salient (as indicated by increased time to respond to the direction of the tilted superimposed bar) when the column was part of a border that had high orientation contrast. However, if the collinear column was not part of the border, or the border had a non-collinear column, response time was shorter (see Figure 2.8a). They also found that when looking at saliency of different cues (colour and luminance), increased saliency by means of colour cues (see Figure 2.8b) did not affect performance, but increased saliency of luminance cues (see Figure 2.8c) facilitated performance. This implies that when computing saliency of stimuli, collinear grouping interacts with orientation, but collinear grouping does not interact with other cues such as luminance and colour to compute saliency.

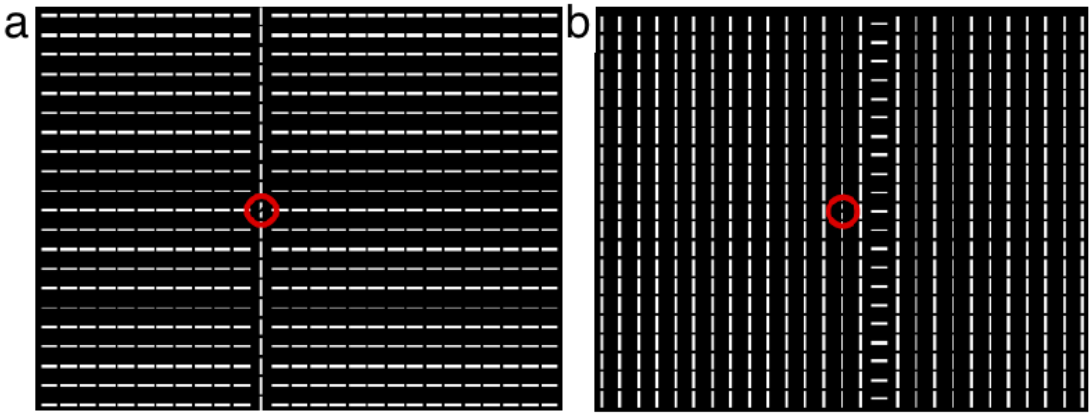


Figure 2.7. The target, which is a superimposed bar, is highlighted with a red circle. The target could appear on a) the salient pop-out column, and b) the non-salient column, with collinear arrangement. From Jingling and Tseng (2013).

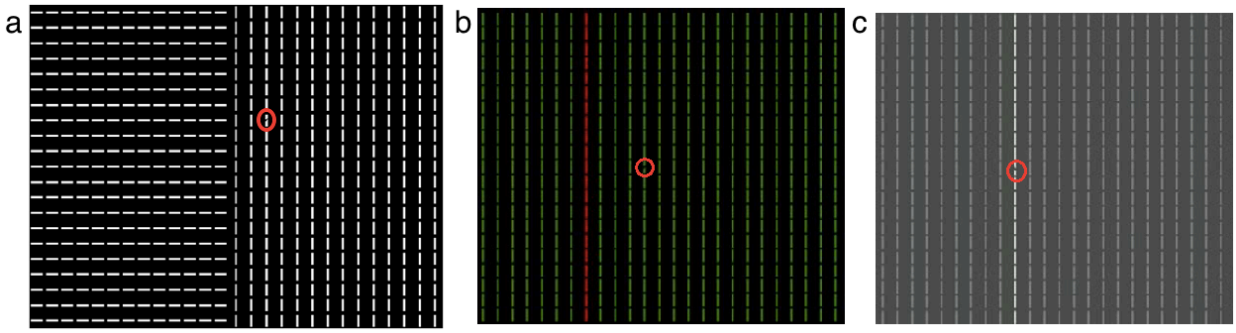


Figure 2.8. The target, which is a superimposed bar, is highlighted with a red circle. a) The target appears on the collinear non-border column. The stimulus was made salient by means of b) colour and c) luminance. The target could appear on the b) non- salient or c) salient column. From Li, Tseng, and Li (2013).

Hunt, Mattingley, and Goodhill, (2012) studied the effect spatial arrangement of edges had on perceptual saliency. They constructed stimuli that had Gabor patches with different arrangements of orientation: co-oriented, co-circular, natural scene, and random (see Figure 2.9). A binocular rivalry paradigm was used, whereby two stimuli were presented simultaneously in different colours (blue and red) for 40 seconds, and participants had to indicate which colour was more dominant. One stimulus was always the random arrangement, while the other stimulus could have been either the co-oriented, co-circular, or naturalistic. They found that the random arrangement was consistently more dominant compared to the other three arrangements. This suggests that random arrangements with high entropy are perceptually more salient.

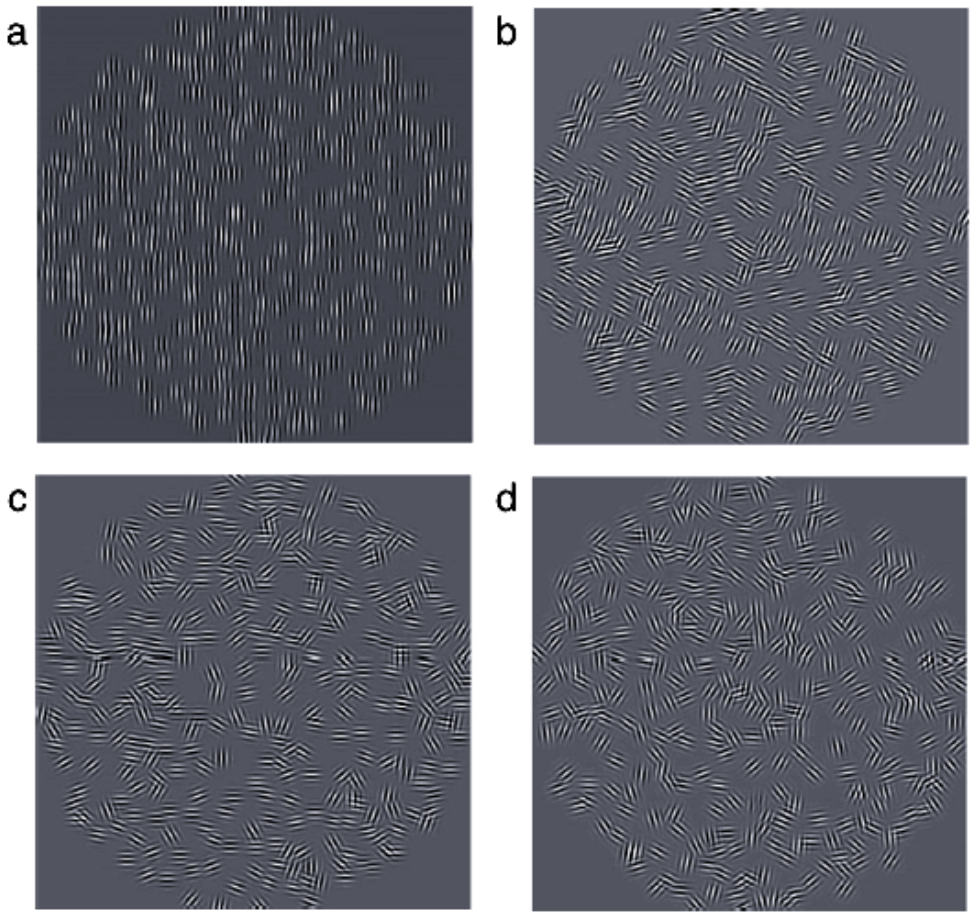


Figure 2.9. The four types of stimuli used by Hunt and colleagues. a) co-oriented, b) co-circular, c) naturalistic, and d) random arrangements. From Hunt, Mattingley, and Goodhill, (2012).

Furthermore, Vancleef et al., (2013) studied the influence that element arrangement had on both discrimination and segregation tasks. They found that the arrangement of elements, either random or regular (see Figure 2.10a and b), did not influence performance in the discrimination task, but regularly positioned elements had facilitated performance in a segregation task, in which participants had to indicate if the edge was straight or curved (see Figure 2.10c and d). This indicates that spatial arrangement plays a role in shape discrimination, as needed by a segregation task, but not in a discrimination task in which discriminating two regions from each other will suffice.

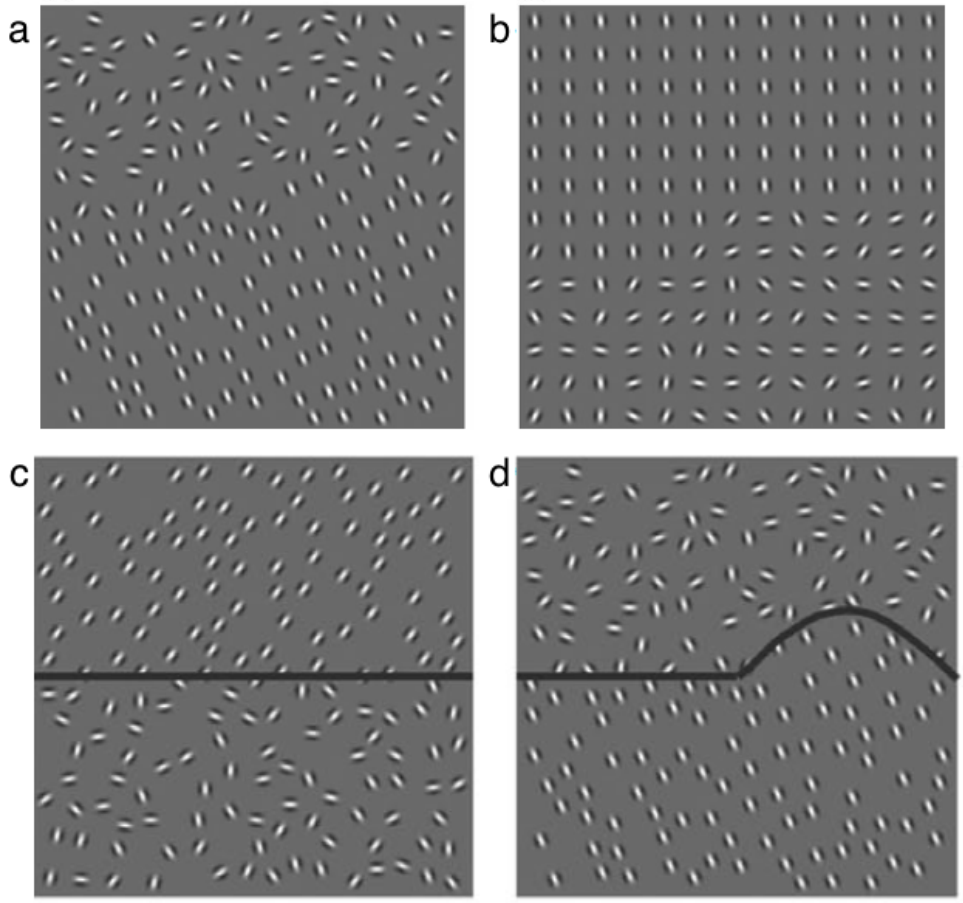


Figure 2.10. a) Stimulus with random positions, b) Stimulus with regular positions, c) Stimulus with a straight edge, d) Stimulus with a curved edge. From Vancleef et al., (2013).

All of these taken together suggest that there is more to texture segregation than just orientation contrast at a boundary region. The configuration and arrangement of elements within a texture region have been shown to facilitate the segregation process. Two possible ways to disrupt these configural effects would be to 1) increase the orientation jitter, and 2) randomly position the elements. In the case of the former, increased orientation jitter would result in a greater variation in orientation of the line elements, which would reduce/remove the occurrence collinearity. Likewise, when the arrangement of the elements are randomised i.e. not uniform, collinearity will be disrupted. In Experiment 1 and 2, we specifically performed control checks using these manipulations to ensure that the findings we got were not confounded by configural effects.

2.4 Evidence for Edge-based and Region-based Mechanisms

Does our visual system have different mechanisms for texture analysis? Based on the literature, the two mechanisms that are used to analyse textures are the edge-based mechanism, and region-based mechanism. The edge-based mechanism detects discontinuities within a texture image, and segmentation occurs due to the explicit boundary that is formed, while the region-based mechanism groups neighbouring elements that are similar together, and as a result, an implicit border is formed between the coherent regions. In the literature, some findings suggest that these mechanisms operate exclusively (e.g. Landy & Bergen, 1991), while others believe that they function together (e.g. Wolfson & Landy, 1998). Considering how one of the primary goals of this thesis is to investigate the operation of these mechanisms in texture segregation, this section will discuss several studies that support the presence of either an edge or region-based mechanism.

The structure gradient of a stimulus is the change in orientation ($\Delta\theta$) over the distance in which the change in orientation occurs (Δx). Therefore, a stimulus that has an orientation change occurring over a larger distance would not have a clearly defined edge boundary, but a rapid change in orientation (over a short distance) would produce a clearly defined edge boundary. Nothdurft, (1985b) found that with increasing Δx, the $\Delta\theta$ also has to be greater to preserve performance levels.

Landy and Bergen (1991) studied the effect the structure gradient had on texture segregation. In their study, participants were shown a stimulus for 33ms, and had to identify which part of a square had one of the four corners diagonally cut off (See Figure 2.11). Like Nothdurft (1985b), they found that performance declines with decreased orientation contrast, $\Delta\theta$, and increased distance over which orientation change occurs, Δx. It is important to note that the structure gradient relies on orientation change occurring over a distance, and in this particular

study, identification of the missing corner occurred over a short distance, and this could have influenced the outcome of the study. They also found that performance declined with lower spatial frequencies. Additionally, performance was independent of viewing distance, which suggests that it was the spatial frequencies of the object, as opposed to retinal spatial frequency, that was driving the performance of the task. They propose a model that is described in Section 2.7.

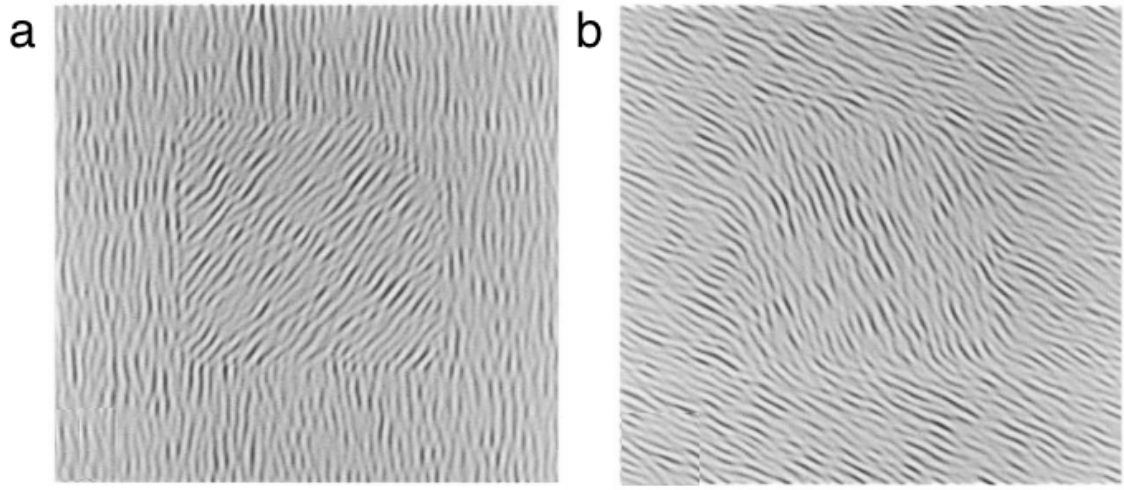


Figure 2.11. a) Orientation change $\Delta\theta = 36°$ and distance over which orientation change occurs $\Delta x = 0°$, b) Orientation change $\Delta\theta = 36°$ and distance over which orientation change occurs $\Delta x = 2.5°$. From Landy and Bergen (1991).

Wolfson and Landy (1998) also examined the importance of edge and region-based mechanisms in texture segregation using orientated line segments. There were two presentations of stimuli – one in which both regions of the texture had the same mean and standard deviation of line elements, and the other where the mean or standard deviation of line elements was different between the regions. Participants were tasked with identifying the interval that contained texture regions with different means or standard deviation. They believed that different mechanisms would be employed based on the information present in the stimuli. They achieved this by having stimulus segregation based on either orientation

difference (Figure 2.12c and d), or orientation variability (Figure 2.12a and b). For orientation variability, orientation variation between each region would vary, and segmentation can only occur by gathering information from the overall regions. 75% threshold values were obtained, and they found that when regions were defined by orientation contrast, abutting textures had improved performance compared to separated textures (orientation thresholds for Figure 2.12c were lower compared to orientation thresholds for Figure 2.12d). When the regions were segregated based on orientation variation, abutting and separated textures had similar performance. Thus, when the texture was segregated by orientation contrast, the information for segregation was most prominent in the edges, and when the edges were separated, performance was impaired. However, when the texture was segregated by orientation variance, the information required for segregation was not just at the edges, but the entire region. As a result, having the two regions abutted or separated did not play a role in the segregation process for orientation variance. Thus, in the absence of orientation contrast between the regions, region-based mechanisms can aid in texture segregation. Taken together, these results indicate that both an edge-based and region-based mechanism exists, and the mechanism used for the segregation process depends on the information present in the stimuli.

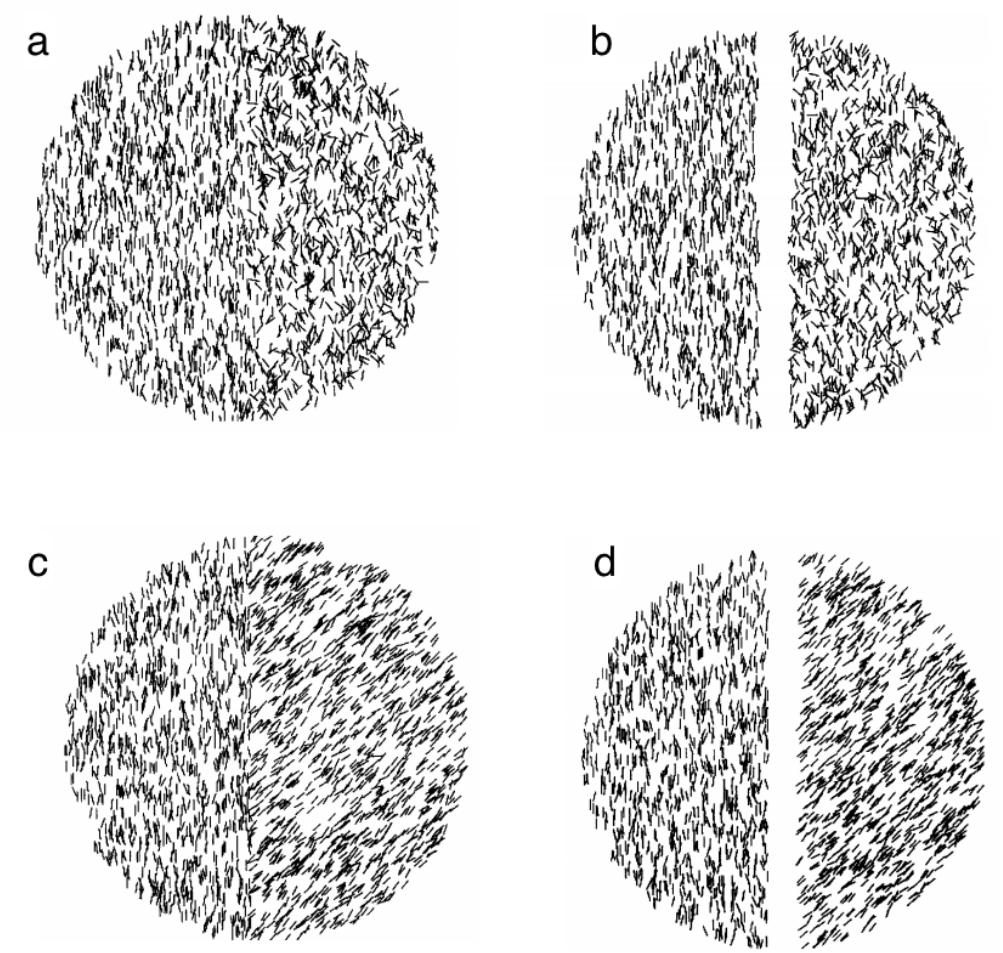


Figure 2.12. a) The texture region on the left has a 10° SD of texture elements, while the region on the right has a 30° SD of texture elements. The two regions are abutting. b) The texture region on the left has a 10° SD of texture elements, while the region on the right has a 30° SD of texture elements. The two regions are separated. c) The texture region on the left has elements with mean orientation 0°, while the region on the right has elements with mean orientation of 45°. The two regions are abutting. d) The texture region on the left has elements with mean orientation 0°, while the region on the right has elements with mean orientation of 45°. The two regions are separated. From Wolfson and Landy (1998).

However, Gurnsey and Laundry (1992) found that texture regions that were abutting (Figure 2.13a) or separated (Figure 2.13b) did not affect performance in a segregation task. It is possible that this difference in results could be due to the different types of stimulus used (see Figure 2.13), as the line elements relied on orientation difference at the boundary for segregation, while the Xs and Ls relied on feature differentiation.

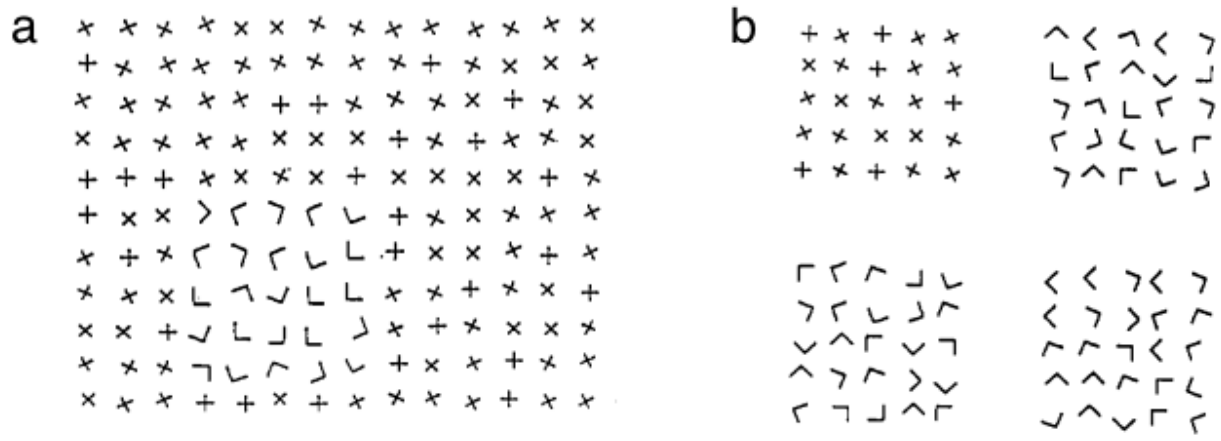


Figure 2.13. a) A smaller patch embedded within the larger figure in the abutting condition. b) Regions separated with white space for the separated condition. From Gurnsey and Laundry (1992).

Further to this, Norman, Heywood, and Kentridge (2011) had participants perform an odd-one-out task, in which they were shown a stimulus for 5 seconds (Figure 2.14), and then had to indicate if the upper or bottom part of the 3 regions was the odd-one-out. Like Wolfson and Landy (1998), they found that performance for orientation contrast was better when the 3 regions were abutting as opposed to when they were separated, and there was no difference in performance for orientation variability. However, for the orientation variation condition, they found that participants were on average faster by 124ms to respond to the separated textures than the abutting textures. They postulate that this is due to the ease with which a region-based mechanism can perform an analysis when the regions are separated as they have clearly defined regions of interest.

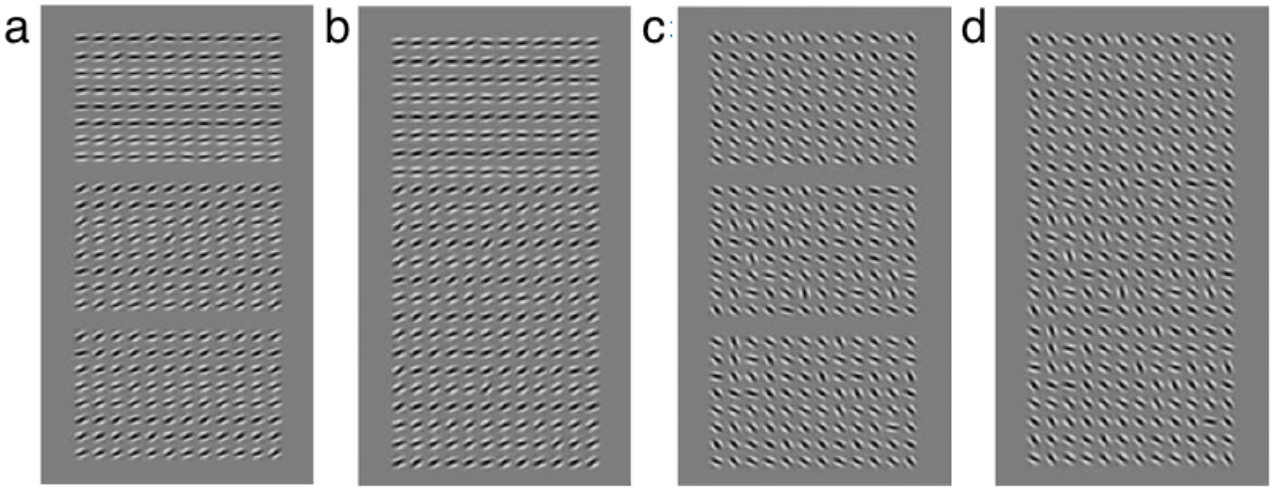


Figure 2.14. a) 3 separated regions of texture stimuli, with the top region having a difference-in-mean orientation. b) Same as A, but in an abutting condition. c) 3 separated regions of texture stimuli, with the top region having a difference-in-variance orientation. d) Same as C, but in an abutting condition. From Norman, Heywood, and Kentridge (2011).

Based on the literature, there is support for both a region and edge-based mechanism during segregation. However, whether these mechanisms work in parallel, or are independent of each other remains to be known. This is something that we will explore further in the thesis. That is, we have designed experiments to explicitly test whether texture segregation involves an edge or region-based mechanism. However, as noticeable from the studies discussed here, most of these stimuli comprise of textures with abrupt edges i.e. a point at which there is an extreme discontinuity of stimulus property. These abrupt edges are only one type of texture variation, and as we will discuss in the following section, texture stimuli can also have smooth variations. Therefore, our intentions are not only to investigate whether texture segregation is mediated by either an edge or region-based mechanism, but to also investigate if either of these mechanisms can account for **both** smooth and abrupt texture variations (see Section 2.5 below for details).

2.5 Smooth and Abrupt Texture Variations

In the literature about texture segregation, there is a consensus that textures segregate easily when there is an abrupt texture gradient with high orientation contrast (e.g. Landy & Bergen, 1991; Wolfson & Landy, 1998). However, segregation performances for textures that do not have an abrupt texture gradient (smoothly varying orientation) are more diverse. In this section, we will discuss studies that have used stimuli with smooth texture variations. Our interest in smooth texture variations stems from our aim to investigate the mechanisms of texture segregation i.e. edge or region-mechanisms (see Section 2.4). A smoothly varying texture would not have an abrupt texture edge, which would make an edge-based mechanism ineffective in segregating the regions of the texture. We therefore believe that to investigate the mechanisms of texture segregation, both smooth and abrupt texture stimuli should be used. This is the foundation of all the experiments (both psychophysics and eye tracking) conducted in this thesis, whereby we used different types of stimuli that either had an abrupt or smooth texture variation.

First, as previously discussed, the greater the orientation difference at a texture boundary, the easier the segmentation process. In general, segregation performance is better when a texture edge is defined by an abrupt change in texture contrast, and performance declines when the orientation change occurs over a greater distance (e.g. Nothdurft 1985a; Landy & Bergen, 1991). However, there is evidence of texture segregation in the presence of smoothly varying change in orientation contrast (i.e. without an abrupt texture gradient). Gurnsey and Laundry (1992) found that 'blurring' (smoothed texture boundary) did reduce performance in a discrimination task, but this decline in performance was less than 10%, and therefore not very important. Furthermore, when they looked at textures that had abutting regions, or separated regions, they found that performance was not affected by the layout. Thus, an analysis of

region was being performed to make the discrimination as the separated region stimuli had no boundary information to make the decision. While this does not negate the role of boundary detection for scene segmentation, it does suggest that texture segregation is not carried out solely by texture gradients, but that regional properties of the texture are influencing the outcome.

However, the stimulus used by Gurnsey and Laundry (1992) for the discrimination task consisted of Xs and Ls. Other researchers have also used smoothly varying stimuli, but with orientation change (Kingdom, Keeble, & Moulden, 1995). A two-interval forced-choice task was used, in which participants had to correctly identify which interval had stimulus with modulated orientation. The other interval always consisted of orientation with zero modulation. Orientation modulation in this case was defined by amplitude values, in which the orientation change throughout a single cycle of orientation modulation was equal to the amplitude value. The orientation modulation varied as a function of either a square-wave (SQ), or sine-wave (SN), which respectively have an abrupt texture gradient, and smoothly varying texture gradient (See Figure 2.15a and12b). They found that performance was better for the SQ condition compared to the SN condition.

Kingdom and Keeble (1996) subsequently included a missing fundamental (MF) condition, which is obtained by removing the fundamental harmonic component from a square-wave. The resulting waveform thus produces an orientation change which has both an abrupt and smooth texture gradient (see Figure 2.15c). As before, they found that SQ had better performance then SN, and MF had overall the lowest performance. When threshold values were obtained, and the sensitivity to the orientation modulation plotted for each waveform according to increasing spatial frequencies (increased number of cycles), the SQ and SN stimulus had a similar pattern to their response, despite the fact that SQ produced better

performance overall, but the MF displayed a steeper decline with increased spatial frequencies. They propose that there is a common linear mechanism that detects both smooth and abrupt changes in the texture. Their model is discussed in Section 2.7.

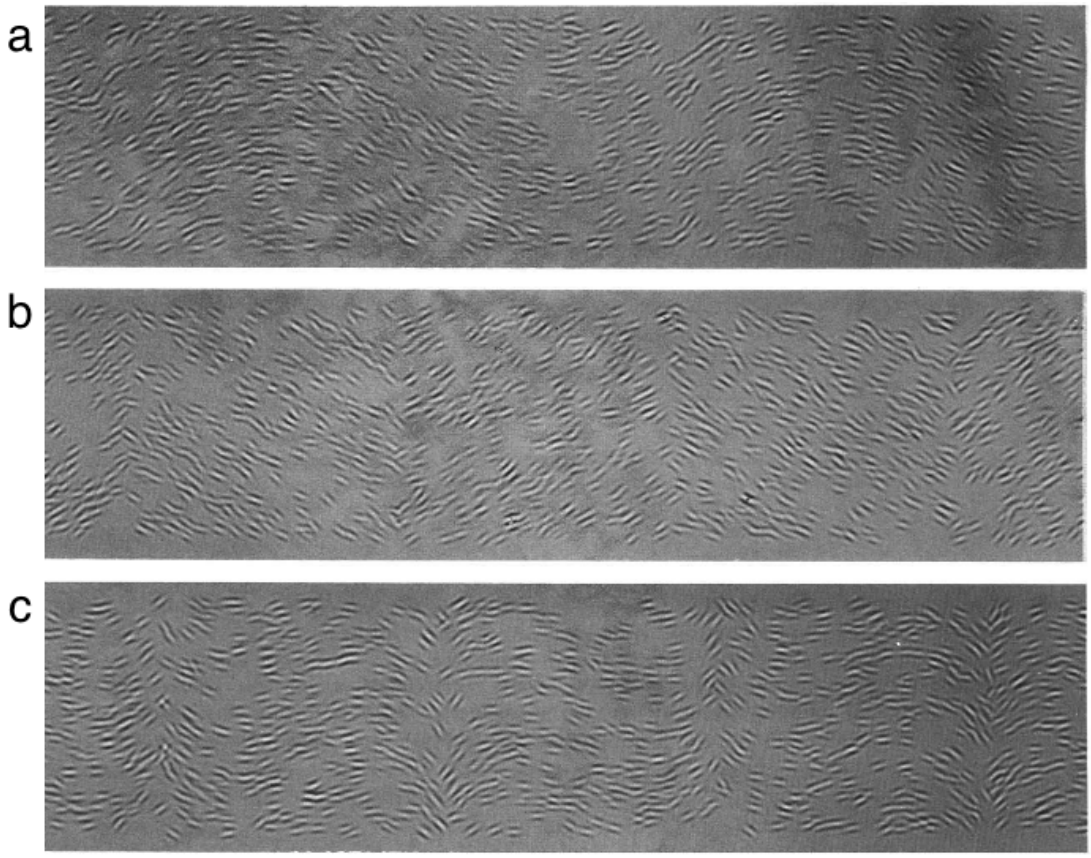


Figure 2.15. The a) Sine-wave, b) Square-wave, and c) missing-fundamental waveform that defines the change of orientation of Gabor patches. From Kingdom and Keeble (1996).

Ben-Shahar and Zucker (2004) proposed that two curvatures (tangential and normal curvatures) are present in texture patches, and found that texture segmentation is influenced strongly by disruptions to these curvatures. They propose that at each point, vector fields that are tangent (E_T) and normal (E_N) to the point are created. The tangent and normal vector field representation of that point is then used to create a frame at each point, and a frame field representation in the image domain (see Figure 2.16d), which consists of information of the tangential and normal curvatures. The frame field representation at point q as it moves along

direction V (blue arrow) results in a representation of the initial rate of change in the tangential ($\nabla_V E_T$) and normal ($\nabla_V E_N$) direction, which then produces two values K_T and K_N which respectively are the tangential and normal curvatures. To study the role these curvatures play in segregation, they manipulated the discontinuities in the tangential (ΔK_T) and normal curvatures (ΔK_N). There were 3 possible fundamental behaviours of the curvatures, which were $\Delta K_T = 0$, $\Delta K_N = 0$, and $\Delta K_T = \Delta K_N$. These respectively indicated a maximal discontinuity in the normal curvature, and no discontinuity in the tangential curve; maximal discontinuity in the tangential curvature, and no discontinuity in the tangential curve; and equal discontinuity in tangential and normal curvature (Figure 2.16 a-c). They tested the effect these curvatures had on performance of tasks independent of the $\Delta\theta$ between (orientation change between 2 regions) and $\Delta\theta$ within (orientation change within a single region). In accordance with Nothdurft (1991), they found that performance in the segregation task was better with larger $\Delta\theta$ between and smaller $\Delta\theta$ within. Interestingly, when taking the different curvatures into account, segmentation was better for $\Delta K_T = \Delta K_N$ at increased values of $\Delta\theta$ between, while performance was at chance level for $\Delta K_N = 0$ and $\Delta K_T = 0$. Thus their findings suggest that texture segregation is not only influenced by texture gradients, but also discontinuities in texture curvatures, especially when the discontinuity is equalised, and maximised simultaneously, for both the normal and tangential curvatures ($\Delta K_T = \Delta K_N$).

The idea that curvatures within an image can influence segregation is certainly interesting. In fact, Bell, Gheorghiu, Hess, and Kingdom (2011) found that the perception of shapes involves an (in their case, luminance-polarity selective) intermediate-level of processing that integrates orientation and position information across space to compute curvatures prior to encoding of global shapes. That is, the mechanism that detects shapes is responding to curvatures that were calculated **from** the local information about orientation and position.

Thus, it is a possibility that a mechanism that segregates regions of a texture also responds to the curvatures present in an image.

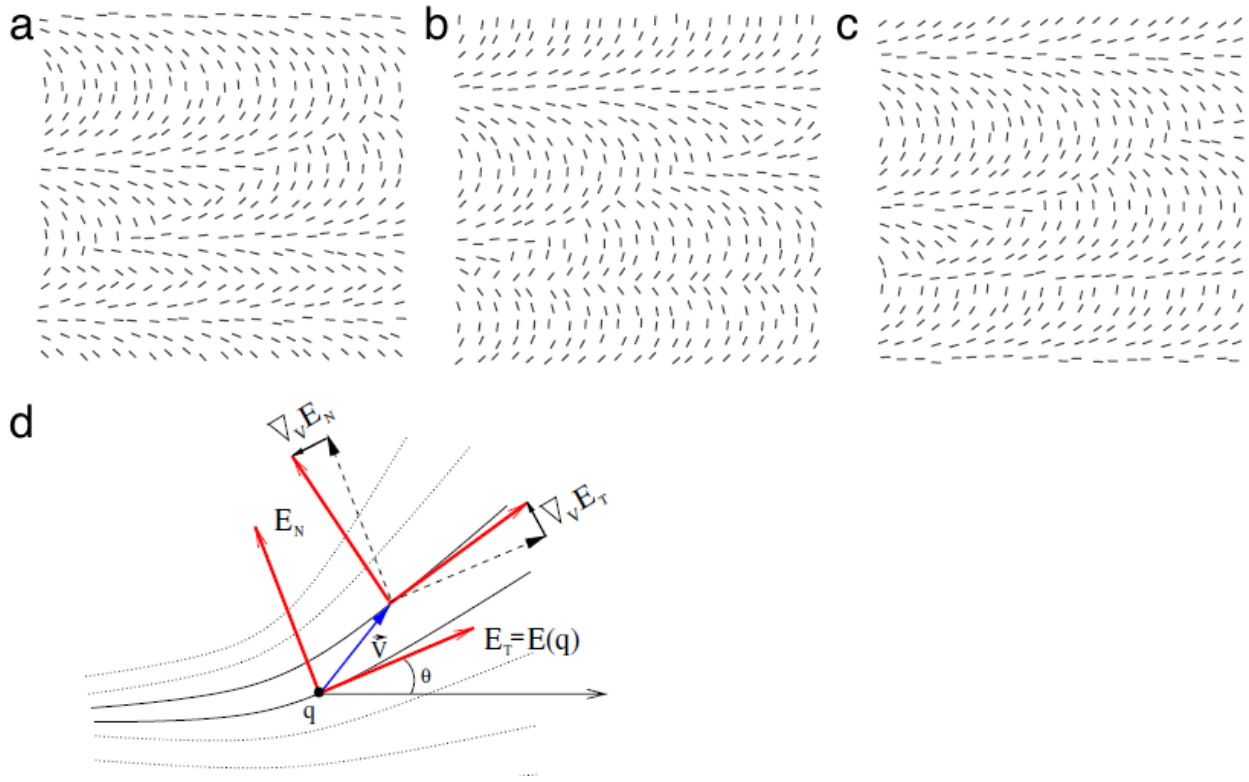


Figure 2.16. Left pointing triangle stimulus used in the experiment. For all 3 examples, the Δθ between = 20° and Δθ within = 90°. The curvature discontinuities a) ΔKT=0, b) ΔKN=0, and c) ΔKT=ΔKN. d) A representation of the frame field of an orientation defined texture. From Ben-Shahar and Zucker (2004).

Subsequently, Ben-Shahar (2006) found evidence of perceptual singularities (i.e. boundaries between perceptually coherent regions) with smoothly varying texture (see Figure 2.17). In a simple tracing task, and discrimination task, it was found that even with their smoothly varying stimulus that kept the orientation gradient constant, participants were able to reliably segregate regions. While texture gradients were unable to reliably produce texture segregation, perceptual singularities were identified based on discontinuities in the tangential and normal curvatures. However, a key point to note about their study is the use of line integral convolution (Carbal & Leedom, 1993) textures, as opposed to line segments, which

are more commonly used. It has been found that boundaries that abut have increased salience compared to discontinuous line elements (Sturzel & Spillman, 2001). This could possibly account for the perceptual singularities found in their study. A proposed model is discussed in Section 2.7.

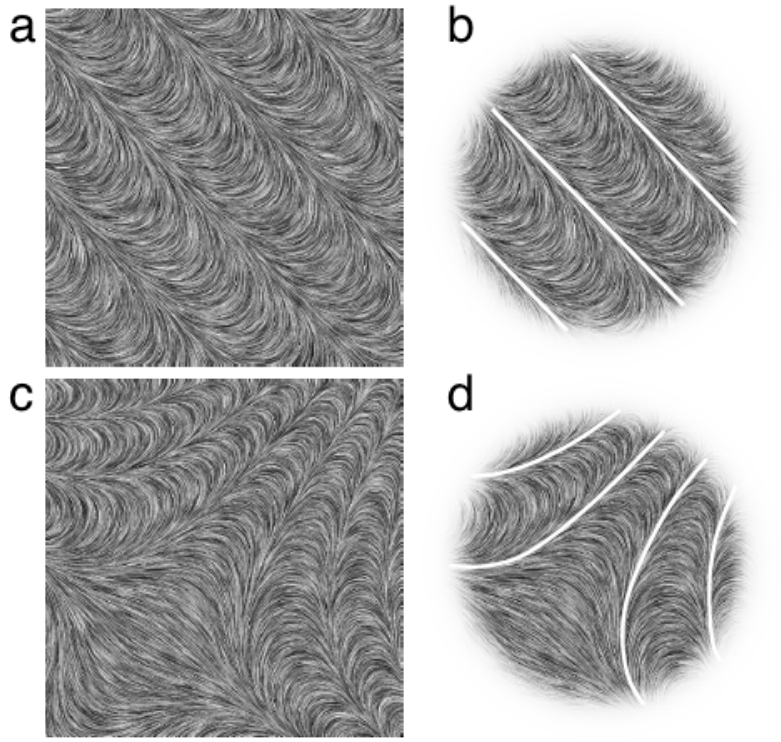


Figure 2.17. a & c) Example of stimuli used in Ben-Shahar's study, and b & d) the perceptual singularities in white lines. The orientation change that defines the perceptual singularities are small. From Ben-Shahar (2006).

To summarise, a substantial number of studies have focused on texture segmentation with abrupt edges, claiming that a rapid change in orientation is required for texture segregation. However, some studies have also shown that texture segregation is possible when there is a gradual change in orientation. In some cases, there seems to be a moderate decline in performance when smooth-edged textures are used, whereas others find no difference in performance between smooth and abrupt texture edges.

2.6 Filling-in Hypothesis

There has been psychophysical (e.g. Caelli, 1985) and physiological (e.g. Lamme, Rodriguez-Rodriguez, & Spekreijse, 1999) evidence in the literature that scene segmentation occurs due to region classification. This is aided by the fact that surface segregation occurs via a filling-in mechanism. This was also discussed in Section 2.2 (Neurophysiology of Texture Perception), where single-cell recording and EEG techniques were informative in describing the time course of neuronal activation across the border and surface area of the figure. This section will cover psychophysical studies that used the masking paradigm to investigate the filling-in mechanism. This is of particular importance to us as we have conducted two experiments (Experiments 1 and 2), the aims of which were to investigate the presence of a filling-in mechanism with texture stimuli that have abrupt and smooth texture variations.

This filling-in phenomenon was first explored by Paradiso and Nakayama (1991) using luminance cues. Evidence of surface filling-in was found, as the presence of a mask appearing on the target did inhibit the spread/filling-in as long as the mask was present at the location before the perceptual spread reached that area. It was also found that the spread occurred in a gradual manner, as increasing the distance between the edge and the mask resulted in longer durations before the spread was inhibited.

Caputo (1998) also found evidence of surface filling-in, but unlike Paradiso and Nakayama (1991) who found a masking effect that started at 0ms, the masking effect only started at about 40ms in the study. Participants were presented with a texture pattern in their periphery that had a 90° orientation contrast and were required to perform an adjustment task to adjust the luminance of the line elements within the masking square to match the luminance of the line elements outside of the masking square. Considering that both the studies were exploring

different cues, it is possible that this difference in masking effect was due to different mechanisms of filling-in for luminance and texture cues.

In a second experiment, Caputo (1998) used texture patterns with varying degrees of orientation contrast (0°, 9°, 18°) and had participants view two intervals of texture stimuli, a texture-mask interval and a comparison texture patch interval, and perform a comparison procedure to regulate the luminance of the comparison patch to match the luminance of line elements within the mask of the texture-mask interval (see Figure 2.18). It was found that the masking effect was reduced as the orientation contrast increased. This was hypothesised to occur due to the spread throughout the entire texture stimulus when a segregated border was absent. However, in the presence of a texture border, the spread begins from the texture border resulting in a decreased masking effect. Caputo defined two stages of region filling-in, where stage one spreads isotropically the mean stimulus luminance i.e. independent of orientation of the elements. This first stage is an early process that operates from 0 – 40ms, and allows for surface representations of discontinuous elements. The second stage employs an anisotropic spread that extracts the border at around 40 – 80ms, followed by region filling-in through lateral excitatory connections that peaks at about 120ms.

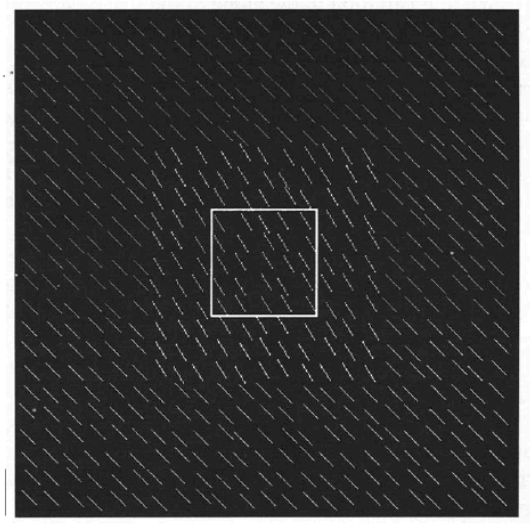


Figure 2.18. Example stimuli used by Caputo (1998). The centre of the texture has an 11x11 segmented patch with an orientation contrast of 9° compared to the background elements. The masking stimulus (white square) is shown here for illustrative purposes only, in the experiment, the masking square is presented after the texture stimulus is presented.

Sturzel and Spillman (2001) also found evidence of a 2 stage filling-in process. They explored how saliency of stimulus affected fading time of texture stimuli, and found that increased saliency (by increasing orientation contrast at the border) resulted in stronger resistance to fading. Fading time was the time between stimulus onset and participants' first response that the target in their periphery disappeared. The first response was used, as eye-movements and blinks typically caused the target to reappear. They used two types of orientation contrast; one with an abutting border (continuous lines) and the other without (short segmented lines), see Figure 2.19. When the border was abutting, there was an increase in time for the stimulus to fade compared to when the border was not abutting. When the border was abutting, more time was needed to breakdown the border (cancellation of border), but when the stimulus border was spatially non-contiguous (short lines), a gradual substitution would have sufficed without explicit cancellation of the border. This is in line with Caputo (1998) who suggested 2 mechanisms for border cancellation and substitution from the surrounds.

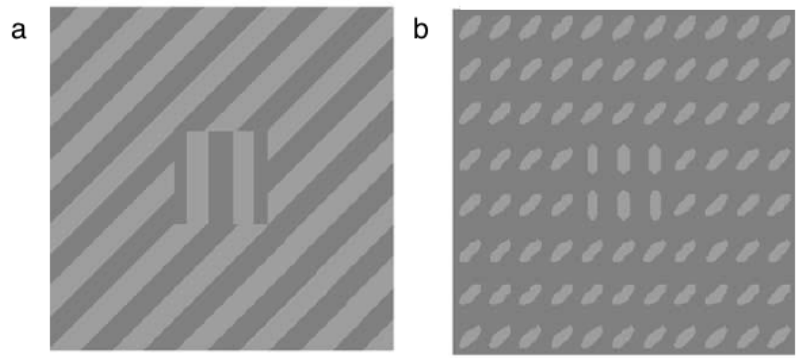


Figure 2.19. Texture stimuli used as defined by orientation contrast of a) a square wave grating and b) short lines. Varying degrees of orientation contrast were used (5°, 10°, 15°, 20°, 30°, 45°, & 60°). From Sturzel and Spillman (2001).

Attar et al., (2007) found a similar effect of fading time and salient stimuli. They used texture stimuli that had uniformly oriented line elements in the centre, and random line orientation in the backgrounds, or vice versa (Figure 2.20a and b respectively). Magnitude estimation ratings were used for participants to rate the perceived strength of the stimuli from 1 to 9. Stimuli with a randomly oriented figure and uniform background were found to be highly salient compared to stimuli with uniform centre and random background (Figure 2.20b was perceived as more salient then Figure 2.20a). Accordingly, texture contrasts that are perceived to be stronger took longer for the central figure to fade (i.e. the central region looked to be the same as the surround). This lent support to the notion that the perceived saliency of a texture boundary did influence the surface filling-in. It also suggests that randomly oriented lines in the surround have a stronger surround suppression (inhibition of neuronal responses due to stimulus outside of the receptive field) compared to uniformly oriented lines in the surround.

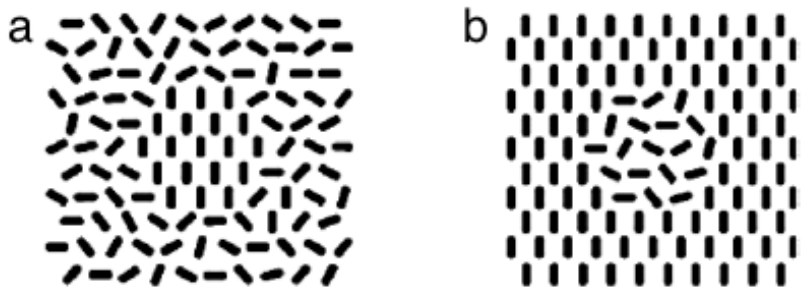


Figure 2.20. a) The texture comprised randomly oriented lines in the surround, and uniformly oriented lines in the centre. b) The texture comprised uniformly oriented lines in the surround, and randomly oriented lines in the centre. From Attar et al., (2007)

Motoyoshi (1999) also found similar filling-in evidence, with the masking effect occurring at -150 to 150ms. Participants were shown a 25×25 grid of horizontal and vertical line elements for 17.7ms, and it was found that highly salient stimuli could still be perceived in the masked area (region that was not perceived due to the mask interrupting the spread). In their case, highly salient referred to low concentrations of one type of elements compared to another type (e.g. low concentration: texture comprise of only 2% horizontal lines, while the remaining 98% are vertical lines). For that example, any horizontal lines within the masked area will still be perceived, but the vertical lines will not. This implies that a mechanism for rapid pattern segregation occurs independently and before a region filling-in mechanism. They extended the findings to apply a region filling-in mechanism that operates for brightness, colour, orientation, and spatial frequency stimuli.

A strong argument for the filling-in process is the presence of a spreading mechanism. As discussed above, in the presence of a boundary contour that is defined by an orientation contrast, the spreading mechanism originates from the border present in the texture, and moves inwards (Caputo, 1998). Su, He, and Ooi (2011) found that when scaled according to cortical distance, the spreading speed remained at a constant rate (28.7cm/s). The spreading speed was measured by the perceived size of a horizontal bar. Participants had to gauge the

size of a horizontal bar, with increased spreading resulting in the perception of a longer horizontal bar. However, when comparing the speed of the spread, stimulus that is presented monoptically spread faster compared to stimulus that is presented dichoptically. It has also been found that the spread time can be influenced by small eye-movements (Yokota & Yokota, 2007). However, further exploration found that it was the small but fast eye-movements (80-200Hz), not slow eye-movements (10-50Hz), that were strongly influencing the spread time (Yokota & Yokota, 2010).

To note, all of the studies discussed here have used variations of the masking paradigm to investigate filling-in. Likewise, the studies discussed in Section 2.2 only used neurophysiological techniques. Moreover, the stimuli used typically had abrupt texture variations – without considering the effects of stimuli with smooth texture variations. In the experiments conducted as part of this thesis, we ensured the use of both types of texture variations, and used these different types of stimuli to investigate this filling-in mechanism (see experiments presented in Chapter 4).

2.7 Models of Texture Segregation

In this section, we describe computational models of texture segregation performance. These models typically operate with the stimulus as the initial input of the model, and the predicted experimental outcomes as the output. The in-between stages of the model are concerned with the operations that are involved in early visual processing. As such, these models are extremely useful in understanding the process/mechanisms involved in texture segregation. Of the three broad types of models we discuss here, the Filter-Rectify-Filter model is the most relevant, as we have come to find that this model fits best with our data. The model we have created also fits well into this Filter-Rectify-Filter framework.

2.7.1 Filter-Rectify-Filter model

A band pass linear spatial filter (i.e. a filter that allows frequencies within a certain range to pass through, while attenuating frequencies outside this range), akin to that of a cortical simple cell, will be able to detect an edge/border if a stimulus was defined by a difference in luminance, wherein there will be a peak response at the edge of dark-light stimulus. However, for texture-defined stimulus (e.g. Chapter 1, Figure 1.1), the average luminance for the figure and ground region is the same, and therefore a mechanism as such cannot account for extraction of texture-defined edges. A more suitable model for texture-defined edges would be a model based on filters selective to specific spatial frequencies and orientations (channels tuned to specific spatial frequency and orientation). A model that employs such a filter is the Filter-Rectify-Filter (FRF) model. Different researchers (e.g. Caelli, 1985; Malik & Perona, 1990; Landy & Bergen, 1991) have used variations of the FRF model to explain texture segregation.

The FRF model of texture perception has a basic 3-stage design (see Figure 2.21) that is commonly used to explain texture segregation (Landy 2013; Landy & Graham, 2004). The 3 stages generally work as follows;

1) First stage linear filtering: this stage models cells in the primary visual cortex that apply an orientation-selective linear spatial filter to the input.

2) Second stage non-linearity: a non-linearity is applied to the output of the first stage filters. This can be achieved by squaring of the filter response by full-wave rectification, or half-wave rectification.

3) Third stage linear filtering: large-scale filters are applied to the rectified output. Edge extraction is modelled in this stage as the linear filters are orientation-tuned and bandpass, which would reveal texture boundaries.

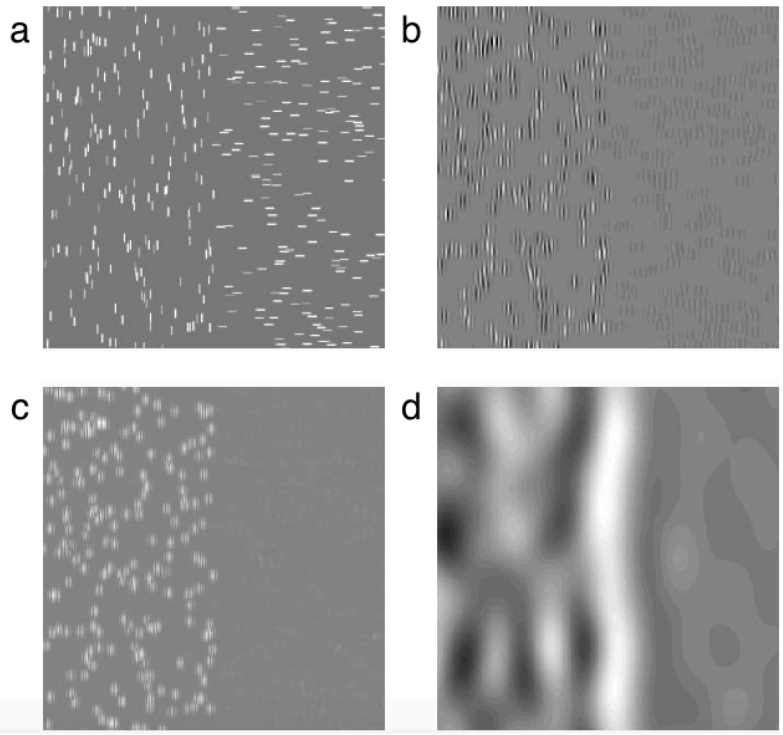


Figure 2.21. a) Example texture stimuli used as an input. b) First stage vertically tuned orientation filter is applied to image A. The left-side vertical lines display larger responses than the right-side horizontal lines, but overall, both regions are strong, resulting in identical outputs on both sides of the image. c) Squaring of the output of the second stage results in stronger responses in the left-side vertical field. d) In the third stage, a second spatial filter that is vertically tuned is applied, which results in a peak vertical response at the edge. From Landy and Graham (2004).

Bergen and Adelson (1986, 1988) proposed a model that segments textures surfaces based on the extraction of simple properties, such as the low-level mechanisms that are tuned to size, from linear filters. This mechanism that is tuned to sizes can be constructed simply by applying a linear centre surround receptive field, which is then full-wave rectified. They suggest a two stage cascade of local energy. In the first stage (linear filtering followed by the

full-wave non-linear rectification), primary energy measures are acquired by spatial averaging. In the second stage, the secondary energy measures are acquired by a set of linear filters. This second stage is what defines the location of the texture boundary. By their own accord, this is not a full model of the texture segregation. However, it does posit that texture perception is possible with a filtering operator that works on image intensities. In their example, they tuned the mechanisms to be sensitive to size, but theoretically, it could be used for any other simple properties like colour, and so forth.

Landy and Bergen (1991) (also Bergen and Landy (1991)) proposed a FRF model that simulates texture segregation on the basis of local oriented energy. The model initially analyses an input via spatial filters that are tuned to specific spatial frequencies and orientation. The resulting output generates four images; vertical (V), horizontal (H), diagonal down-and-left (L), and-right (R). The local energies of the four images are derived by a square-wave non-linearity, which is then averaged across a small region to achieve a pooling response.

Finally, the differences in opponent signals (orientation discrimination) are calculated from their orthogonal responses (i.e. H^2-V^2 and R^2-L^2), and contrast normalisation process occurs to allow for texture energies to be computed independent of variation contrast. Edge extraction occurs when there is a rapid change in opponent signals (see Figure 2.22). Thus, regions with increased orientation contrast ($\Delta\theta$), and regions where the change of orientation occurs over a shorter distance (Δx) will experience a quick shift in opponent signals, leading to edge extraction.

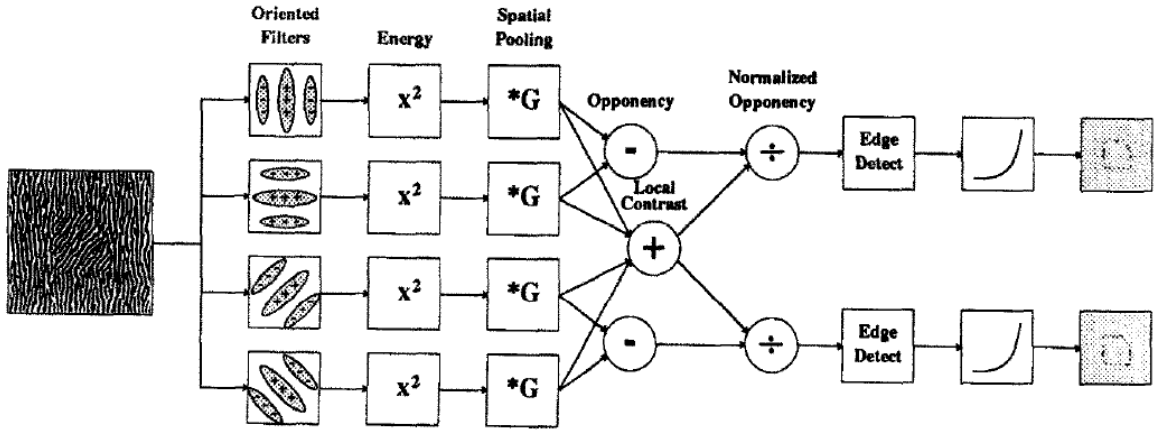


Figure 2.22. Texture analysis model by Landy and Bergen (1991).

Unlike the model based on the extraction of edges from simple properties (Bergen & Adelson, 1986, 1988), the model proposed by Landy and Bergen (1991) is tuned to local energy derived based on orientation signals. This model is invariant to other features like luminance contrast and, size of the stimuli, while still extracting edges from the texture surface.

The model by Landy and Bergen (1991) was able to fit the data from their psychophysical experiments, in which performance declined with decreased orientation contrast ($\Delta\theta$), and increased distance in which the orientation change occurred (Δx). However, increasing Δx does not affect the model's performance as much as the behavioural data. Decreasing $\Delta\theta$ does strongly decline performance in the modeling data, which fits with the behavioural data.

The two models described thus far follow the framework of the 3-stage FRF model. However, both models posit that texture segregation occurs due to localising gradients. In contrast to this, Caelli's model that also follows the framework of the 3-stage FRF model, proposed that segmentation occurs due to region classification (Caelli, 1985). The model initially generates an image that the segmentation process can act upon. This is done first by a

set of eight spatially oriented bandpass linear filters, which then goes through a compressive non-linearity and rectification process.

The segmentation process occurs in 3 stages: impletion/filling-in, correlation, and grouping. The impletion stage is when excitation transfers/spreads from a highly responsive region, to a less responsive region. The correlation stage is when a consensus is formed from the set of eight filters at each location, and the grouping stage clusters responses that belong to the same region. As a result, the output highlights surfaces/regions first, which is then followed by an implicit border being formed between the regions. A key point of this model is that it can account for the fact that texture segregation occurs even without an abrupt change in orientation (e.g. Gurnsey & Laundry, 1992).

A model that can account for the extraction of edges based not only on abrupt changes in texture gradient (localising gradients) but also from smooth changes in texture gradients (region classification) was proposed by Kingdom and Keeble (1996). The 3 classes of linear models proposed are; single-channel peak amplitude, most detectable harmonic, and single-channel r.m.s. (root mean square).

The single-channels peak amplitude model posits that the detection of the orientation modulation is achieved by a single linear operator that has a peak response. On the other hand, the most detectable harmonic model posits that for detection to occur, the amplitude of the stimuli has to be the same as the detection threshold for its most detectable harmonic reach (sensitivity to the most detectable harmonic present). Finally, the single-channel r.m.s. model posits that detection occurs due to the r.m.s. responses of a single linear operator.

Psychophysical results were analysed according to these three models, and it was found that the single-channel r.m.s. had the best account for the psychophysical results, in which abrupt

changes in texture orientation resulted in better performance compared to smooth changes in orientation, and textures with both smooth and abrupt changes had the lowest performance. The single-channel peak amplitude model and the most detectable harmonic model also accounted for the psychophysical data, but to a lesser extent.

2.7.2 Biological models

So far, the models discussed were based on the FRF model framework. The following models discussed are based on models of the Visual Cortex – primarily V1, V2 and V4. Poort et al., (2012) put forth a model that combines the iso-orientation inhibition of boundary detection with the iso-orientation excitation of region filling-in into a single hierarchal network. This is akin to Kingdom and Keeble (1996) who proposed a single linear channel to detect both abrupt and smooth changes in orientation, but in the current model, border detection and surface filling-in are facilitated by area V1 and V4 respectively.

According to their model, a boundary is detected via iso-orientation suppression by a local centre-surround receptive field, where there is strong suppression at regions with homogenous orientations, and weak suppression at orientation discontinuities. This results in a border being formed in the V1m (modelled area of V1) and V2m (modelled area of V2). Their model also uses feedback connections from the V4m (which has a courses neural map due to larger RF) that excites neurons in V1m and V2m that has the same orientation, which results in region filling-in.

The attentional effects are modelled by weaker neuronal responses in V4m (modelled area of V4) when attention is not explicit. As such, feedback connections from the V4m to the V2m and V1m will also be weaker, causing only partial filling-in of the region. When attention is explicit, V4m responses are stronger, and so would the responses to V2m and V1m.

Theilscher and Neumann (2003, 2005, 2007) put forth a model that relies on feedback connections from the V4, like Poort and colleagues, but instead of a feedback input facilitating region filling-in, the bidirectional links within V1, V2, and V4 serves to reduce noise and provides a robust border extraction.

The modeling of each cell in each stage in the hierarchy is achieved by:

1) Initial activation due to bottom-up activity of the RF

2) Feedback from higher areas modulates the initial activation

3) These are sensitive to position and orientation, and as such, cell activity in the lower areas of the hierarchy are modulated by the higher areas

4) Contrast enhancement and normalisation of activation levels occurs via intra-areal centre-surround computation by top-down modulation

5) Steps 1-4 are repeated until a stable activation pattern is formed over the whole model

The model cells in the early visual areas signal local features such as orientation, while mid-level cells are sensitive to texture borders. As a result, V1 and V2 model cells signal position and orientation discontinuities, which are passed along to the V4 model cells. The V4 model cells are thus only reacting to the orientation discontinuities without being influenced by the homogenous regions. When feedback from V4 to V2 and V1 occurs, it will have focused on the discontinuities, which gives the model its resistance to noise data.

Finally, the model proposed by Li (1999a, 1999b, 2000, 2002a, 2002b, 2003) focuses only on the V1 region of the primary visual cortex. The V1 model proposed mimics layers 2-3 of the orientation selective cells, and also takes into account the contextual effects of cells via

horizontal intra-cortical connections. By taking into account contextual influences, the model is able to detect boundaries between regions, even if there are several locations in which orientation contrast is present. This is managed by the model detecting the location where there is a break in input homogeneity (Li, 1999a). The model also accommodates collinear facilitation, iso-orientation suppression (Li, 1999b), general surround suppression, and contour enhancement (Li, 2002a, b) via horizontal intra-cortical connections. This is modelled as the pyramidal cells, which can excite each other monosynaptically or inhibit responses disynaptically by projections to relevant inhibitory neurons. Furthermore, lateral connections are modelled to link cells that have similar orientations, which result in orientation specific contextual influences (Li, 2003).

As a result, the model reliably predicts psychophysical results such as increased saliency to increased orientation contrast, border regions with parallel texture elements, and decreased saliency with increased jitter (Li, 1999a). The model also accounts for the border effect, in which increased saliency is given to the border due to the weaker iso-orientation suppression, and cross-orientation facilitation, in which the figure has high saliency if the orientation of the figure is orthogonal to the background. Theilscher and Neumann (2007) suggest that the V1 model is limited by the spatial integration of information by V1. Compared to behavioural results in which performance decreased gradually with decreasing density of texture elements (Nothdurft, 1985b), the V1 model (Li, 2000) demonstrated a steep decline in performance for spacing elements that were further apart. Theilscher and Neumann's model mimics a gradual decrease in performance akin to that of the behavioural performance. According to them, feedback from V4, which has large receptive fields and will have a broader representation of the contextual information, provides additional information that is similar to the human visual system.

2.7.3 Curvature model

Ben-Yosef and Ben-Shahar (2008) proposed three models where curvature estimations are computed via: spatial RF summation, multi-scale RF interactions, and non-linear gating. These are computational models that are also based on biologically plausible mechanisms, however their model calculates curvatures of textures to detect perceptual singularities (boundaries between perceptually coherent regions) in textures with smoothly varying orientations.

For all 3 models, the computations of perceptual singularities rely on the normal and tangential curvatures. They are also based on several plausible biological mechanisms. One of these mechanisms is the local estimation for the orientation texture, which is done by an even-symmetric receptive field for each spatial location. The second mechanism involves the simple cells in V1 that have odd-symmetric receptive fields to generate estimations of directional derivatives that are perpendicular in direction to the preferred orientation. Finally, using filter convolution and thresholding, the local maxima points are calculated to obtain the resulting Perceptual Singularity Measure (PSM), which is the model predicted results of perceptual singularities.

The 3 models they propose have different computation for the calculations of the maxima points. The spatial RF summation model calculates curvatures based on summation of the responses in the normal and tangential direction. Orientations that have the highest summation would be selected, resulting in a representative map that signals the PSM ridges. The multi-scale RF interaction model employs intercolumnar interaction to calculate curvature values. Here, facilitation occurs for interactions of RFs or filters with similar orientations due to even-symmetric filters. The key difference between these 2 models is that the multi-scale RF interaction model uses filter responses to calculate the degree of

orientation change, while the spatial RF summation model uses summation responses. The filter responses give increased sensitivity to the curvatures, which is better at detecting perceptual singularities.

Finally, the non-linear gating model computes curvatures by a feed-forward three-layer circuit. Unlike the other 2 models, which calculates the curvatures estimation of each location in isolation, this model uses channel interactions in the early processes to produce a dominant orientation map. In this model, both even and odd-symmetric receptive fields are used to produce derivatives which are then used to calculate the dominant orientation. From this, the tangent and normal curvatures are calculated while all other maps are 'shut down' via a non-linear gating mechanism.

Comparing model performance, the non-linear gating model detects perceptual singularities best. The authors also found that their model is able to account for boundary contour detection of textures segregated by an abrupt change in orientation, but typical FRF models cannot detect contours from their smoothly varying stimuli.

The model thus uses calculations of curvatures to estimate the location of perceptual singularities, which implies the role of curvatures in segregation. This is in accordance with the Ben-Shahar and Zucker (2004) who found that equalizing and maximizing the tangential and normal curvatures of line elements increased participants performance in a segregation task. Even when smoothly varying stimuli with a constant orientation gradient was displayed, participants were reliably able to trace regions of perceptual singularities based on curvature discontinuity (Ben-Shahar, 2006).

2.8 Spatial Vision and Aging

Visual abilities are typically impaired during the course of normal i.e. non-pathological aging. Some of these impairments are linked to the deterioration of the optical properties of the eye, which cause visual deficits in low-level visual functions, such as visual acuity and contrast sensitivity (Owsley & Burton, 1991). However, optical properties alone do not account for these age-related declines in human performance, as the physiological properties of visual neurons also partially contribute to the diminished visual abilities (Schmolesky, Wang, Pu & Laventhal, 2000; Betts, Taylor, Sekuler, & Bennett, 2005).

For one of the experiments that will be presented in this thesis, we investigated the effects of healthy aging in texture segregation. Our rationale for doing this will be described in greater detail in Chapter 6, but briefly, we had the unique opportunity to investigate this using the different types of stimuli i.e. texture stimuli with either smooth or abrupt texture edges, and also the model we created. That is, since our stimuli and model were created to investigate the early stages of visual processing, we were able to study the effectiveness of visual processing of older adults compared to younger adults. This experiment to investigate texture segregation in aging also had the added benefit of further testing our model, and consequently testing what parameters of a texture stimulus creates a change in the outcome of the model (these were also the objectives of the experiments presented in Chapter 5). This section will first cover the fundamentals of spatial vision, which would lay the framework for understanding the spatial frequency manipulation we used in the aforementioned experiment. Finally, we will review studies that explore both the affected and spared perceptual abilities of older adults.

2.8.1 Spatial Vision

In 1822, Baron Jean Fourier discovered that any waveform (as a function of time, space, or other variable) could be broken down into its constituent parts of sine and cosine waves of specified frequencies, amplitude, phase, and orientation. Fourier analysis is a method in which an image can be broken down (or analysed) into a summed set of sinusoidal gratings (periodic patterns in which luminance varies across space) that differ in the four aforementioned parameters. Conversely, the inverse Fourier transform synthesises an image from a set of sinusoidal gratings (DeValois & DeValois, 1990). Figure 2.23 shows examples of sinusoidal gratings defined by the four parameters.

The spatial frequency is the number of oscillations per unit of distance. In vision, the unit most useful is the visual angle, which is the angle a viewed object subtends at the eye. Therefore, the customary description of spatial frequency is cycles per degree (cpd), where one cycle (oscillation) depicts one pair of dark and light bars. Consequently, a higher spatial frequency will have more number of cycles with narrower bars (see Figure 2.23 a and b). The amplitude of a grating is the difference between the dark and light bars (trough and peak) divided by 2. The contrast of a grating is related to the amplitude, as both are measures of intensities of the waveform. However, the contrast of a visual stimulus is given by the Michelson contrast:

$$\text{Contrast} = \frac{L_{max} - L_{min}}{L_{max} + L_{min}}$$

where L_{max} and L_{min} are respectively the peak and trough of the waveform. The phase of a waveform refers to a position with respect to a reference point (e.g. does a grating begin with a dark or bright bar). The orientation of a grating is the rotation angle (specified in degrees) of the grating from vertical (Palmer 1999; DeValois & DeValois, 1990).

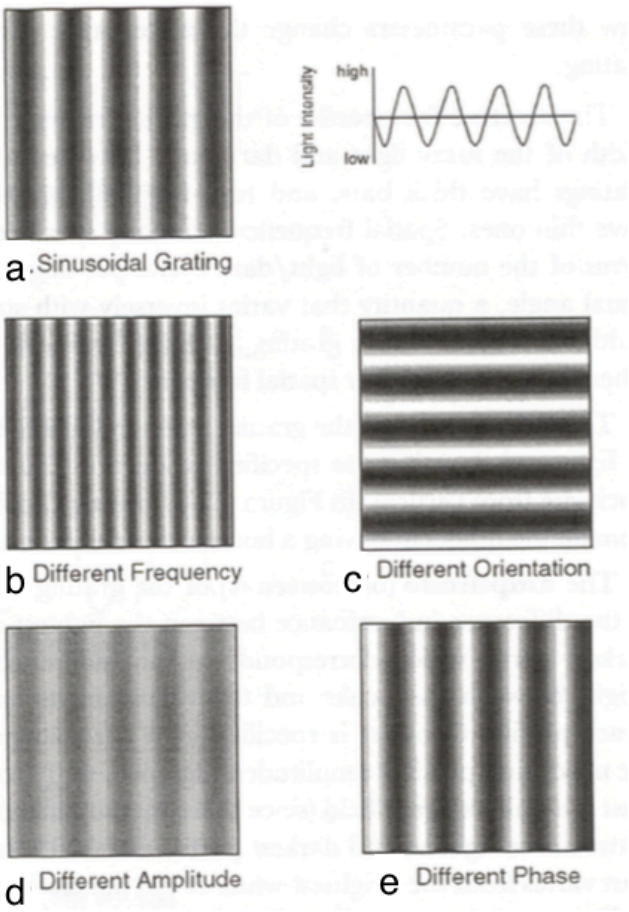


Figure 2.23. Sinusoidal gratings are shown for a) a standard grating with the graph on the right showing luminance varying sinusoidally across space. Four other comparison gratings of, b) higher spatial frequency, c) different orientation, d) different amplitude, and e) different phase is shown. From Palmer (1999).

Fourier had shown mathematically how useful sinusoidal gratings are (since images are made up of these), and in vision research, these are certainly just as useful. However, the spatial extent of the sinusoidal grating will influence visual interpretation. A sine wave with an infinite number of cycles would have extremely precise spatial frequency information (since the filter that is receiving input from an extended sine wave would only receive information about that precise spatial frequency), but with considerable amount of uncertainty in terms of spatial location. In contrast, a sine wave filter of limited cycles would have a broader spatial frequency passband, but with greater certainty of spatial location. Thus, a trade-off is seen

between uncertainty in terms of spatial frequency and spatial location (DeValois & DeValois, 1990). Gabor (1946) found that using a Gaussian fall-off reduced the uncertainty in the spatial frequency or location domain, making them more ideal for visual stimulus. Gabor patches, which are sine waves multiplied by a Gaussian function take the form as shown in Figure 2.24.

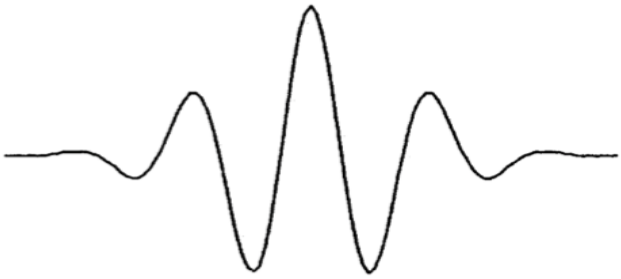


Figure 2.24. The product of a sine wave multiplied by a Gaussian distribution. From DeValois and DeValois (1990).

2.8.2 Multiple Spatial Frequency Channels & Contrast Sensitivity

There is evidence of the existence of multiple spatial frequency channels in vision (Campbell & Robson, 1968; Blakemore & Campbell, 1969). We will discuss this in more detail after a brief explanation of sensitivity.

The sensitivity (1/threshold) of a stimulus is the minimum stimulus intensity required to just about detect a stimulus. In laboratory settings, sensitivity is typically measured as a function of some sort of systematic manipulation. A well-known example is the contrast sensitivity function (CSF), where contrast thresholds are measured as a function of the spatial frequency of a grating (Figure 2.25a).

Blackmore and Campbell (1969) measured each participant's CSF, after which they had the participant adapt to a grating of a particular (fixed) spatial frequency. The CSF was re-

measured again, and the outcome was that the visual system adapted to the grating (therefore becoming less sensitive), but only for the spatial frequencies near the spatial frequency of the adapted grating i.e. there was a drop in sensitivity around the spatial frequency of the adapted grating, but not for spatial frequencies that are higher or lower. The results of the study were explained in terms of spatial frequency channels. That is, the shape of the CSF is actually based on a set of overlapping, narrow-band channels, all of which are sensitive to different range of spatial frequencies (see Figure 2.25b).

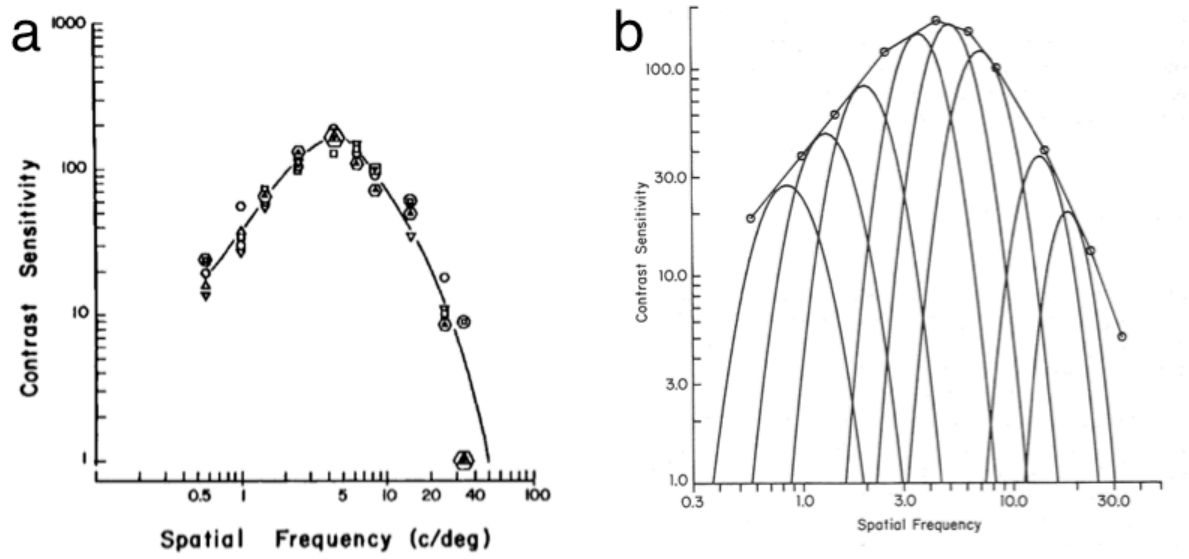


Figure 2.25. a) Contrast sensitivity, which is the inverse of contrast thresholds, plotted as a function of spatial frequency, b) Contrast sensitivity function as the envelope of many narrowly tuned spatial frequency channels. From DeValois and DeValois (1990).

2.8.3 Perception in Aging

While it is fairly common that older age carries greater risks for ocular diseases, healthy aging also impairs visual function for two reasons: optical changes in the eye (e.g. reduced pupil size), and age-related changes to the neural processing of visual stimuli (McKendrick, Chan, & Nguyen, 2018). The latter of which we will be discussing in this section.

Owsley, Sekuler, and Siemsen (1983) found that sensitivity (as measured by the method of adjustment) to static sinusoidal gratings of low spatial frequency (0.5 and 1cpd) remained unaffected with age, but sensitivities decreased for medium to high spatial frequencies (2cpd and above) starting from the ages of 40 – 50. The authors conducted this study as previous studies had found differing results on contrast sensitivity in adulthood – e.g. when testing spatial frequencies between 0.2 and 6.4cpd, Arden (1978) reported no differences in sensitivity between young and old adults, while Skalka (1980) reported that older adults performed worse than younger adults. Owsley and colleagues recruited older adults who were free from any ocular disease that could affect the results, and also ensured that all participants wore their best optical corrections for the test distance (this was done as previous studies had not been explicitly clear about the optical state of the participants). They also tested a large sample (n=91) of participants aged between 19 and 87 years of age so that the data gathered would be representative of the older population (as opposed to just a few older individuals, some of whom might have atypical data).

Schefrin, Tregear, Harvey, and Werner (1999) used sinusoidal gratings with a Gaussian envelope (Gabor patch) to test the contrast sensitivity of older and younger participants at scotopic light levels. The authors used a two-interval forced-choice paradigm to test a range of spatial frequencies between 0.2 and 3cpd. They found that compared to the younger age groups, the older age groups had decreased contrast sensitivity for all spatial frequencies tested, with a greater loss in sensitivity observed for spatial frequencies below 1.2cpd.

The above studies have focused on the effects of aging on contrast sensitivity. The remainder of this section will discuss studies that investigate the effect of aging on visual perception. This is of particular interest as the process of healthy aging differentially affects different

aspects of vision (e.g. Owsley, 2011; Hutchinson, Arena, Allen, & Ledgeway, 2012; McKendrick, Chan, & Nguyen, 2018).

Some studies report that older adults face deficits in tasks requiring high-level (or second-order) processing (Habak & Faubert, 2000). Second-order processing refers to the second-order properties of a stimulus such as contrast, texture, or depth, and first-order properties refer to luminance or colour. In their study, they had old and young participants perform tasks using first-order (luminance defined) or second-order (contrast defined) stimuli. They tested participants in 5 different tasks – 1) discriminate the orientation (is it vertical or horizontal?) of static gratings, 2) discriminate the orientation (is it vertical or horizontal?) of drifting gratings, 3) discriminate the direction of motion (is it moving left or right?) of drifting gratings, 4) & 5) second order gratings are presented in two intervals, and the participants task is to determine if the gratings in interval 1 and 2 had the same orientation (4) and direction of motion (5). Only for task 1 did old and young participants perform equally as well. For tasks 2 and 3, older participants performed slightly worse than younger participants for the first-order stimuli, but performance for older adults was greatly impaired for the second-order stimuli. For task 4 and 5, older adults were once again greatly impaired compared to younger adults. To sum up, compared to younger adults, older adults' performance declined severely with second-order stimuli, and not much for first-order stimuli – suggesting a disassociation between the mechanisms used to process first and second-order stimuli. They argue that this deficit in the elderly can be attributed to processing-related changes, as second-order stimuli require more cortical processing levels.

In fact, Faubert (2002) proposes that the deficits in perceptual processing in the elderly become more apparent when the computation load of the task reaches a high level of complexity. When more complex networks have to be accessed (either because of

simultaneous processing of complex tasks, or perceptually complex stimuli like second-order stimuli) to complete a task, the aging brain does not have the neural resources to perform the task. However, perceptual functions involving minor neural circuitry show no difference in performance because alternative neural networks may be used to compensate, thereby preserving performance. Moreover, Tang and Zhou (2009) conducted a study similar to Habak and Faubert (2000) using a large sample size (n=141, ages=19 – 79), and found similar results. They additionally found that the decline in second-order processing begins earlier (but with a slower rate of progression) compared to the decline in first-order processing. Finally, Bertone, Guy, and Faubert (2011) used Landolt-Cs that were defined by first (luminance) and second-order (contrast) properties, and measured old and young participants' performance to determine the gap positions. They found that for the luminance-defined targets, older adults performance was slightly impaired compared to the younger adults (difference of 2 dB). For the contrast-defined targets, sensitivities were significantly lower for the older adults (difference of 6 dB). Interestingly, despite the findings from the studies mentioned, there are studies (these will be discussed below) that find that older adults do not constantly show impairment in visual perception of second-order stimuli.

For the field of motion perception, Snowden and Kavanagh (2006) claim that studies (e.g. Wood & Bullimore, 1995) typically use random dot patterns (that contain a range of different spatial frequencies) in studies with the elderly. These stimuli are confounded with the loss of sensitivity to high spatial frequencies in the elderly. Therefore, it is not known if elderly participants' poor performance is due to loss of motion sensitivity, or due to partial loss of input of high spatial frequencies. The authors investigated motion sensitivity in the elderly, using sinusoidal gratings (0.5 – 4cpd) that drifted laterally, as well as random dot patterns. Velocity thresholds were found for direction of motion judgements (gratings drifting left or right, and dots moving up or down). They found that older participants performed

consistently worse at all spatial frequencies tested (and also for the random dot pattern), suggesting a deficit in motion sensitivity, and not in contrast sensitivity.

The results above were in terms of the lowest threshold of motion (i.e. the lowest speed of motion in which motion perception is possible), which informs about the processing of motion information at low speeds, but not at high speeds. However, Snowden and Kavanagh were also interested in how older participants would perform at faster speeds of motion. To investigate this, they measured participants' motion coherence thresholds at a range of different speeds (dots moving at 0.5 – 4 degrees per second). They found that older participants' thresholds were higher compared to the younger participants, but only at slow speeds, with fast speeds (≥ 2 degrees per second) showing no difference between both age groups. Therefore, older adults do not necessarily face a general decline in motion perception, just at low speeds.

Moreover, several aspects of motion perception were explored by Allen, Hutchinson, Ledgeway, and Gayle (2010), but with the premise that the deficits seen in motion perception are not due to motion processing as such, but due to the contribution of lower level (first stage of local-motion mechanisms) aspects of visual perception, such as contrast and temporal sensitivity. In the experiment that examined contrast sensitivity, motion coherence thresholds were measured for different contrasts (of dots from background) that varied between 0.004 and 0.33. The authors found that at high contrast, older participants' performance was not affected much compared to the younger participants. However, at lower contrasts, older adults performed markedly worse than younger adults. To examine the effect of temporal sensitivity, motion coherence thresholds were measured at two contrast values (high: 0.34; low: 0.05) for different sampling rates of motion stimuli – 0.28° every 53.3ms and 0.07° every 13.3ms, which results in a smoother perception for the latter. For the younger

adults, the smoother motion stimuli improved performance for both high and low contrast stimuli. On the other hand, smoother motion stimuli improved performance of older adults at low contrast stimuli only. These findings suggest that the encoding of low-level information, especially contrast sensitivity (i.e. not motion integration), drives age-related deficits seen in motion perception.

A surprising finding from Betts, Taylor, Sekuler, and Bennett (2005) is that they have found instances in which old people outperform young people. In their study that also investigates motion perception, they presented participants with 1cpd Gabor patches with gratings that drifted left or right. They manipulated the contrast of the gratings (between 2.8 and 92%) as well as the size of the grating (between 0.7 and 5°), and used a two-interleaved staircase procedure to measure duration thresholds (lowest stimulus duration needed to determine 77% correct response). They found that for low contrast gratings, performance of young and old adults improves with larger sizes. For high contrast gratings, young adults' performance declines with larger sizes, while performance does not vary with size for older adults i.e. only for high contrast stimuli at large grating size, older people perform better than younger people. The authors, as well as Tadin and Blake (2005), propose that this effect was due to age-related reductions in GABA-mediated inhibition, which results in reduced cortical inhibition in the elderly, specifically reduced centre-surround interaction in the visual neurons. That is, when stimulus size is large, greater surround inhibition occurs (as a large portion of the stimuli falls on the surround) causing performance to decline. However, older adults have reduced cortical inhibition, thereby reducing the surround inhibition faced by the larger stimuli, and performance does not decline.

Contrary to the findings by Betts et al., (2005), Hutchinson, Ledgeway, and Allen (2014) found that older adults instead outperform younger adults when the stimulus size is smaller

(<100 deg^2). In their study, motion perception was investigated using random dot kinematograms (same as Allen et al., 2010), while aperture size was varied (they also manipulated the number of elements and the density of the dots in control experiments). They found that motion coherence thresholds increased (decline in performance) with smaller aperture size for young adults. However, for older adults, motion coherence thresholds remained relatively constant with changes in aperture size. In sum, old people outperform young people at small stimuli size. The authors propose that age-related changes in spatial summation (smaller summation areas for global motion processing in older adults), cortical inhibition (reduction in cortical inhibition in older adults), and visual attention (an advantage for old people to have a narrower region of visual space to focus on) could potentially account for the data.

Besides motion perception, the effects of aging in contour integration have also been investigated (Del Viva & Agostini, 2007). Counter integration is measured as the ability to segregate a chain of elements (typically Gabor patches) from a noisy background (Field, Hayes, & Hess, 1993). They manipulated the inter-element distance (which is known to affect performance), and had participants determine which quadrant of the screen had the Gabor chain. Overall, they found that performance declined as inter-element distances increased, and while there was a decline in performance as age increased for all inter-element distance, the most prominent difference between the age groups was for the smallest distances. This is particularly interesting because at smaller distances, the task is easier (as shown by integration sensitivities), and the authors assumed that older people would therefore show less impairment here. To explain these counter-intuitive results, the authors propose the existence of different integration mechanisms, for the different inter-element spacing, that are affected differently with aging (i.e. the mechanism that integrates small contours are not as impaired with age).

Similarly, Hadad (2012) investigated contour integration in the elderly, with the focus of examining the underlying mechanisms of spatial integration. According to Field, Hayes, and Hess (1993), spatial integration of contours depends crucially on collinearity among elements, and less on spatial proximity. That is, when collinearity is high, spatial proximity (elements close together or far apart) does not affect integration. However, when collinearity is low, integration is improved by higher spatial proximity (e.g. Hadad & Kimchi, 2008). Hadad investigated whether this pattern of results changes with adulthood. Participants had to determine the direction an egg-shaped contour was pointing at different levels of collinearity (high and low) and spatial proximity (high and low). It was found that older participants were marginally worse than younger participants, most notably when collinearity was weak. Most importantly however was the interactive effect of collinearity and spatial proximity (where integration is not affected by spatial proximity for highly collinear elements, but when collinearity is low, stronger integration is found for high proximity compared to low proximity contours), which was not affected by age. Hadad claims that while the aging visual system is subjected to increased internal noise that leads to a general reduction in contour integration for older adults, the underlying neural mechanism that spatially integrates information is preserved in old age.

In line with the above, Casco, Robol, Barollo, and Cansino (2011) found that while the integrative mechanisms are not affected by aging, suppressive mechanisms are. To measure these integrative and suppressive mechanisms, the authors measured participants' detectability of deviation from circularity (DFC) for contours with and without background elements. DFC levels were measured as displacement thresholds, which was the amount of displacement needed to identify that an element has been displaced from its position on a closed-contour circle, in a 2-interval forced-choice task. There was also a condition where instead of the contours being formed by tangentially oriented elements, the elements had

alternating tangential and orthogonal orientations. The outcome was that older participants' performance was similar to the younger participants only for the tangential contour with no background, and worse than the younger adults for the other conditions (tangential-orthogonal contour and tangential background). The results are interpreted as older adults being able to integrate the contours as well as younger adults, but performance is impaired in the conditions where they have to suppress the irrelevant information. The authors propose that GABA-mediated reductions in lateral inhibition are what affect the suppressive mechanisms, while sparing the integrative mechanisms.

To focus more on orientation processing throughout adulthood, Delahunt, Hardy, and Werner (2008) investigated the effect of aging on orientation discrimination. The purpose of this was because studies (e.g. Schmolesky, Wang, Pu & Laventhal, 2000) of single-unit recordings of macaque show that neurons in the primary visual cortex have reduced selectivity to orientation in older macaques. Therefore, similar changes are expected to occur in the visual cortex of humans, leading to decreased performance in older adults compared to younger adults. The authors initially measured participants' contrast thresholds (is the Gabor present in the first or second interval?) to Gabors of 1 and 4 cpd. Subsequently, participants' orientation discrimination (is the Gabor tilted left or right of vertical?) was measured for a range of contrast values. The contrast values were based on each individual participant's contrast threshold at five equal logarithmic steps above threshold, as well as maximum contrast. For the final task, orientation tuning was measured using a masking paradigm (contrast detection thresholds were measured for Gabor patches superimposed on sine wave (mask) patterns of differing orientation). The contrast of the mask was set to be 2× the contrast threshold of the participant, and the spatial frequency of the mask was matched to that of the Gabor patch. For the contrast thresholds, they found that contrast thresholds were higher for the older adults, and higher for the 4cpd grating. There was no interactive effect

between age and spatial frequency. For the orientation discrimination and orientation tuning, performance between young and old participants shows no difference when expressed in terms of multiples of contrast threshold, but a difference is found when expressed in terms of physical contrast values (i.e. orientation discrimination thresholds plotted as a function of contrast in 'multiple of contrast threshold' vs. '%'). Furthermore, the bandwidth of the tuning curves (1/2 the height/peak of the tuning curve) showed no difference between both groups of participants. This suggests that age-related increase in orientation discrimination can be explained by contrast sensitivity, and that orientation tuning is preserved in old age.

Govenlock, Taylor, Sekuler, and Bennett (2009) conducted two experiments, the first of which was similar to Delahunt et al., (2008)'s orientation tuning experiment, but without the control of contrast sensitivity. While the shape of the orientation tuning was not exactly identical between young and old adults (older adults reached a lower asymptote before younger adults), the bandwidth of the tuning curves was similar for old and young adults. However, the authors claim that the mask used in the experiment could potentially produce spatial beats that confound the results. To address this, they conducted the experiment again, but using a notched-noise masking (static noise that was broadband in terms of spatial frequency) instead of sine wave masking. They found that the shape and bandwidth of the tuning curve were identical for both young and old participants. This once again shows that orientation tuning is preserved in old age.

Furthermore, Wang, Yu, Fu, Tzvetanov, and Zhou (2018) used a tilt illusion paradigm (where the presence of an orientated stimulus in the surround biases the perceived orientation of the stimulus in the centre) to test old and young adults at two different spatial frequencies – low spatial frequency, which corresponds to a value close to the peak contrast sensitivity of each individual, and high spatial frequency, which corresponds to a value where contrast

sensitivity is lower. There was an overall increase in the tilt illusion for the older adults compared to the younger adults – which is indicative of an increase in lateral inhibition (as a proposed explanation for the tilt illusion is the lateral inhibitions between spatially arranged orientation hypercolumns of neurons in V1 (Georgeson, 1973). They modelled the participants' perception under the assumption that simple feature perception (such as orientation) is derived from the decoding of neuronal activities in V1. Briefly, it is a 2-layer neuronal model, where the first layer of neurons is modelled after simple cells with preferred features of orientation, spatial frequency, and contrast. This is subsequently fed into a second layer that has a connectivity structure with a spatially excitatory centre, and inhibitory surround. The results from the modeling suggest that there is an increase in lateral inhibition with older age (in accordance with their behavioural data), while orientation tuning widths (and thus intracortical inhibition) show no difference in age.

Besides that, Casco, Barollo, Contemori, and Battaglini (2017) studied orientation discrimination in the elderly with the purpose of identifying the efficiency of the encoding (extracting sensory information in the brain) and decoding (readout the extracted sensory information in the brain to form a decision) of information. They measured d' (difference between hit rate and false alarm rate) for a range of orientation offsets (1-12°), wherein Gabor patches were tilted to the left and right of vertical by 1-12°, either in the presence or absence of background/noise elements. The authors hypothesised that a decreased in d' for the older adults at smaller orientation offsets would be indicative of inefficient encoding. On the other hand, if older adults have decreased d' for larger orientation offsets, it is indicative of inefficient decoding (as false alarm rates increase when orientation offset increases). They found that for both noise and no-noise conditions, while younger participants' d' score improved with greater orientation offsets, older participants' d' score did not change with larger orientation offsets. This therefore implied that older adults had difficulties with

decoding as a result of a diminished ability to discard irrelevant information (no-noise: discarding response to non-presented orientation, noise: discarding response to distractors).

Unlike the studies discussed above on orientation which indicate that orientation perception is mostly intact in older age, Kurylo (2006) found impairments in perceptual organisation for the features of orientation and flicker (but not motion or colour) in the elderly compared to younger participants. Grouping Thresholds (the degree of similarity in the stimuli) and Processing Time (duration of stimulus presentation) was measured for the four features. For the feature of orientation, the stimuli comprised '>' and 'V' elements. For both measures of Grouping Thresholds and Processing Time, older participants performed worse compared to younger participants only for the features of orientation and flicker. It is likely older adults do not face a general decline in perceptual organisation, but it is feature dependent, which is associated with the grouping mechanisms specific to each feature.

In sum, the findings of the effects of aging on vision are greatly varied. Within the different domains, there is a trend in which older adults face a decline in visual perception. However, this cannot be generalised to all aspects of vision as there are instances discussed above in which the aging visual system is not compromised, and occasionally has seen benefits.

2.9 Motivation for Psychophysical Experiments

Although many studies have investigated texture segregation using orientation-based stimuli, these figure-ground textures typically have an abutting edge i.e. an abrupt change in texture property to signal the figure border. However, as we have discussed in this chapter, figure-ground segregation is still possible in the absence of a texture edge. Additionally, there is the competing idea of whether the process of texture segregation involves an edge-based or

region-based mechanism. The following psychophysical experiments presented aim to address these issues.

To answer the question of whether a texture edge (i.e. a region where a texture property changes abruptly) is crucial in figure-ground segregation, we constructed figure-ground texture patterns that had different texture variations. The Block profile exemplified a stimulus with an abrupt texture variation, while the Blur profile exemplified a stimulus with a smooth texture variation. We also constructed a Cornsweet profile that had an abrupt texture change at the edge, which tapers off further away from the edge. These figure-ground texture patterns were also used in our eye tracking experiments as the regions in which information of orientation contrast were present were different across the three profiles. However, in those experiments, our main objectives were related to investigating which part of a texture figure was extracted (see Chapters 8 – 10).

In the first experiment presented in Chapter 4, we investigated the role of texture edges in figure-ground segmentation and detection using the aforementioned texture patterns at different viewing durations. This allowed us to address the question of whether texture segregation can be explained by an edge-based or region-based mechanism, and also the involvement of a filling-in mechanism in the segregation process. In the second experiment of Chapter 4, we varied the parameters that controlled the texture variations of the Blur and Cornsweet profiles. The purpose of this experiment was to investigate if there was a common mechanism that is used to segment and detect all texture patterns i.e. not just an edge or region-based mechanism.

We present two experiments in Chapter 5 where the size and aspect ratio of the texture figure is varied. The purpose of these experiments was to determine if the mechanism that segments and detects these orientation-defined textures are fixed in size, or are there different sizes of

this mechanism that changes based on the type of figure-ground textures. In Chapter 6, we piloted a study to investigate segmentation of textures as a function of the spatial frequency of the individual elements that comprise the texture patterns. Subsequently, we present an experiment where we investigate how the aging visual system (i.e. old vs. young adults) segments texture patterns, and how this is affected by the spatial frequency content of the texture pattern. For the final experiment, we investigated the segmentation of texture patterns as a function of viewing distance. The purpose of this experiment was to further investigate the results we found in the previous experiment.

Chapter 3

General Methodology

This chapter will describe the general methodology used in the thesis. Experimental setups that vary from this standard paradigm will be highlighted in the methods section of the various experimental chapters.

3.1 Psychophysical Methods & Psychometric Functions

Asking an observer to report whether or not they were able to perceive a stimulus is the basic procedure of measuring thresholds (Gescheider, 1997). Three of the most common methods are the method of constant stimuli, method of limits, and method of adjustment. The first two of these methods will be discussed in detail below, as they are the methods used in this thesis for the psychophysics experiments. Briefly, the method of adjustment requires the stimulus intensity to be either very high or very low, and the observer will have to respectively decrease and increase the intensity until the stimulus is barely visible. This is very similar to the method of limits, but instead of giving a "yes" / "no" or correct/incorrect response, the participant would adjust the stimulus themselves (Norton, Corliss, & Bailey, 2002).

3.1.1 Method of Constant Stimuli

The method of constant stimuli is the procedure of measuring the performance of a fixed set of stimulus intensities, typically with equal sizes between the different intensities if a linear scale is used. This method randomly draws from the fixed set of levels (usually 5 – 9), which are presented to the observer repeatedly. To compensate for guessing or any sort of response bias, participants are typically forced to make a choice between two or more possible

alternatives, with only one correct response (Norton, Corliss, & Bailey, 2002). The responses for all stimulus intensities are recorded, and a graph of response versus stimulus intensity produces a psychometric function (see section 3.1.3 below). Pilot testing is usually conducted to select the stimulus intensities so that they span the vicinity of the estimated threshold. This is done so that the stimulus intensities chosen produce responses that range between chance (50% for a two-alternative forced-choice procedure) and always correct (100%). The threshold is the stimulus intensity that produces performance at a particular level e.g. 50% "yes" responses, or 75% correct responses (Kingdom & Prins, 2010).

For most of the psychophysical experiments presented in this thesis, the method of constant stimuli was used. For the detection and segmentation tasks, the stimulus intensities in the studies were the orientation contrast between figure and ground of the texture stimuli. For each condition, five magnitudes of orientation contrast were used. These values were determined by piloting the study. For each of the stimulus intensity, 40 different (i.e. non-repeated) trials were generated. Therefore, for experiments using the method of constant stimuli, the threshold calculated from each psychometric function was a result of 200 trials.

3.1.2 Method of Limits

For the method of limits, the intensity of the stimulus is changed based on the prior response of the observer. The most common method of limits is the staircase procedure (or up-down procedure). The first stimulus could have either very high (descending series) or very low (ascending series) intensity, and on each successive presentation, the intensity is changed gradually based on the observer's response (Norton, Corliss, & Bailey, 2002). For example, if a 1-up-1-down ascending series was used, the intensity of the stimulus will start low and will increase with every incorrect response made. Once a participant makes a correct response, the sequence will be reversed and the intensity will decrease. This increase and decrease in

intensities will continue until several reversals have occurred. Thresholds are typically taken as the average of the last few reversals. A variation to this procedure is the interleaved-staircase procedure that uses both an ascending and descending sequence, which have low and high starting intensities respectively. A benefit of this procedure over the single staircase procedure is that participants would not be able to predict the stimulus based on the staircase sequence (Gescheider, 1997).

For this thesis, the QUEST type (Watson & Pelli, 1983) interleaved-staircase procedure was used for the psychophysical experiments that employed the method of limits. The QUEST, which assumes the underlying function to be a Weibull, selects the stimulus intensity for a subsequent trial based on the responses collected thus far. That is, after every trial, the fit of the psychometric function is updated with the newly acquired response. A distinctive feature of the QUEST is that it takes into account prior knowledge about a condition and its estimated threshold. As such, prior to experimental testing, the experimenter would posit the threshold and standard deviation of the various parameters/conditions. In the earlier portion of the experimental run, this knowledge can be used to curb excessive step-sizes by making selections based on the pre-defined values. However, in the later parts of the experimental run, this prior knowledge will play a smaller role, as the newly acquired data will drive the selection of stimulus intensity (Kingdom & Prins, 2010). For the experiments presented in this thesis that used the interleaved-staircase procedure, thresholds for each condition were generated from 120 different (i.e. non-repeated) trials, 60 from each staircase series.

3.1.3 Psychometric Function

A psychometric function is a plot where the y-axis shows the response on a particular psychophysical task (e.g. percentage correct, proportion “yes” response), and the x-axis is a characteristic measure of stimulus intensity (e.g. luminance, size). Figure 3.1 illustrates an

example of a psychometric function. The data points (black dots) on the graph represent the observer's response to different luminance contrasts, which was then fit with a Logistic function. In this example, the threshold is the stimulus contrast (x-axis, dashed line) at which the performance of the observer was correct 75% of the time.

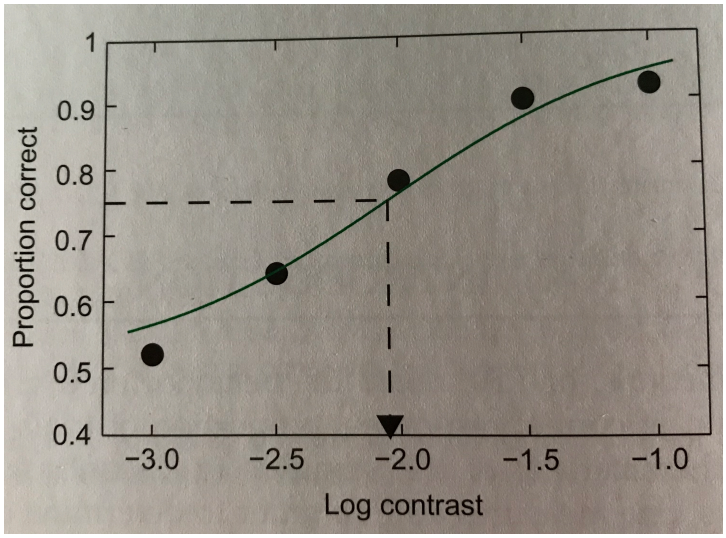


Figure 3.1. Example of a psychometric function taken from Kingdom and Prins (2010).

3.1.4 Curve Fitting

A crucial aspect of the psychometric function is the curve fits used to model the psychometric data. For most of the experiments presented in this thesis, the Weibull function is used to fit the psychometric functions. However, in one experiment related to the eye tracking studies, the Logistic function is used. The equations for the Weibull (F_w) and Logistic (F_l) functions are as below:

$$F_w = 1 - 0.5e^{\left[-\left(\frac{x}{\alpha}\right)^{\beta}\right]}$$

$$F_l = \frac{1}{1 + e^{\frac{-(x-\alpha)}{\beta}}}$$

where x is the value on the x-axis of the psychometric function, and α (alpha) and β (beta) are the parameters of the curve fit. The β (beta) defines the slope/gradient of the curve fit, and the α (alpha) is the threshold when the y-axis corresponds to 50% (Logistic) and 82% (Weibull) correct. The functions were fitted to the psychometric data using a maximum likelihood method, and the error associated with the fit is the 67% confidence interval, which is the error bar on each threshold data point. Threshold values that were not physically meaningful were not included. In the case of our orientation contrast, physically meaningful values are below 90°, as orientation contrast cannot be above 90°.

3.2 Display

Unless stated otherwise in the methods section of the experimental chapters, the display setup used for the psychophysics experiments is as described in this section. The stimuli were generated using PsychoPy 1.83.04 (Pierce, 2007) on an Apple Mac Mini, and displayed on a Formac ProNitron 21/750 CRT monitor that had been gamma corrected. The luminance of the background was 44.8 cd/m^2. The screen resolution was set at 1024×768, and had a refresh rate of 120Hz. At a viewing distance of 133cm, the pixel resolution was 1 pixel per arc min.

3.3 Stimuli

The texture stimuli comprised anti-aliased white line elements on a grey background that produced a figure-ground texture. The anti-aliasing in PsychoPy was achieved by using a mask on each of the individual line elements, which creates an anti-aliased gradient ¼ the width of the line. The length of each line element was 10 pixels long and 2 pixels wide, with an average space of 15 pixels between the centres of the line elements when the elements were positioned in a grid pattern. The orientation of the line elements could form three different types of orientation profile – Block, Blur, and Cornsweet (see Figure 3.2 for a

schematic representation). The dashed line in Figure 3.2 represents the location of the border (or edge) of the figure. The region beyond the border is the ground, while the region within the border is the figure.

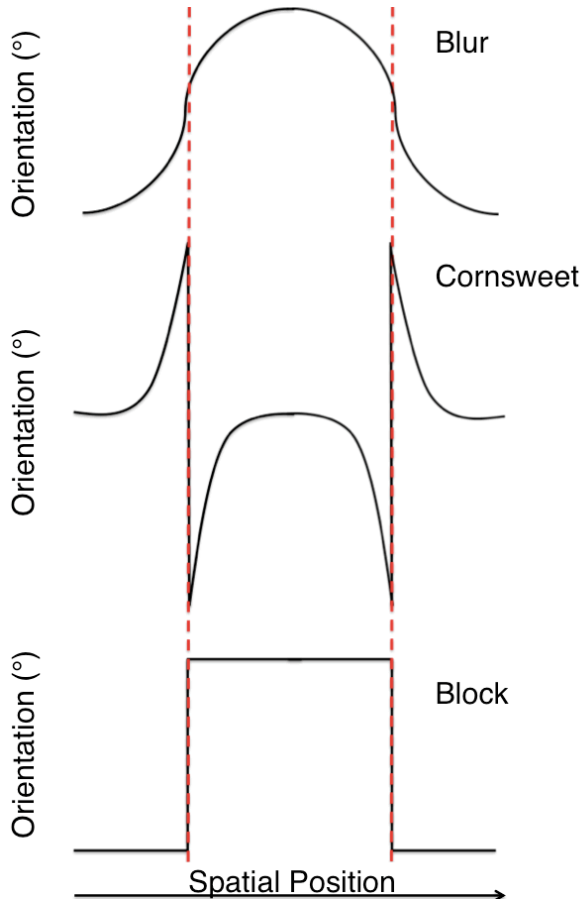


Figure 3.2. A schematic representation of the orientation change as a function of space for the three different orientation profiles used.

For the Block profile, the border of the figure defined the orientation change of the line elements. The orientation of the line elements of the figure region (within the border) was different from that of the ground region by the value of the orientation contrast. The mean orientation of the line elements for the Blur (Equation 1) and Cornsweet (Equation 2 & 3) profiles is defined as follows:

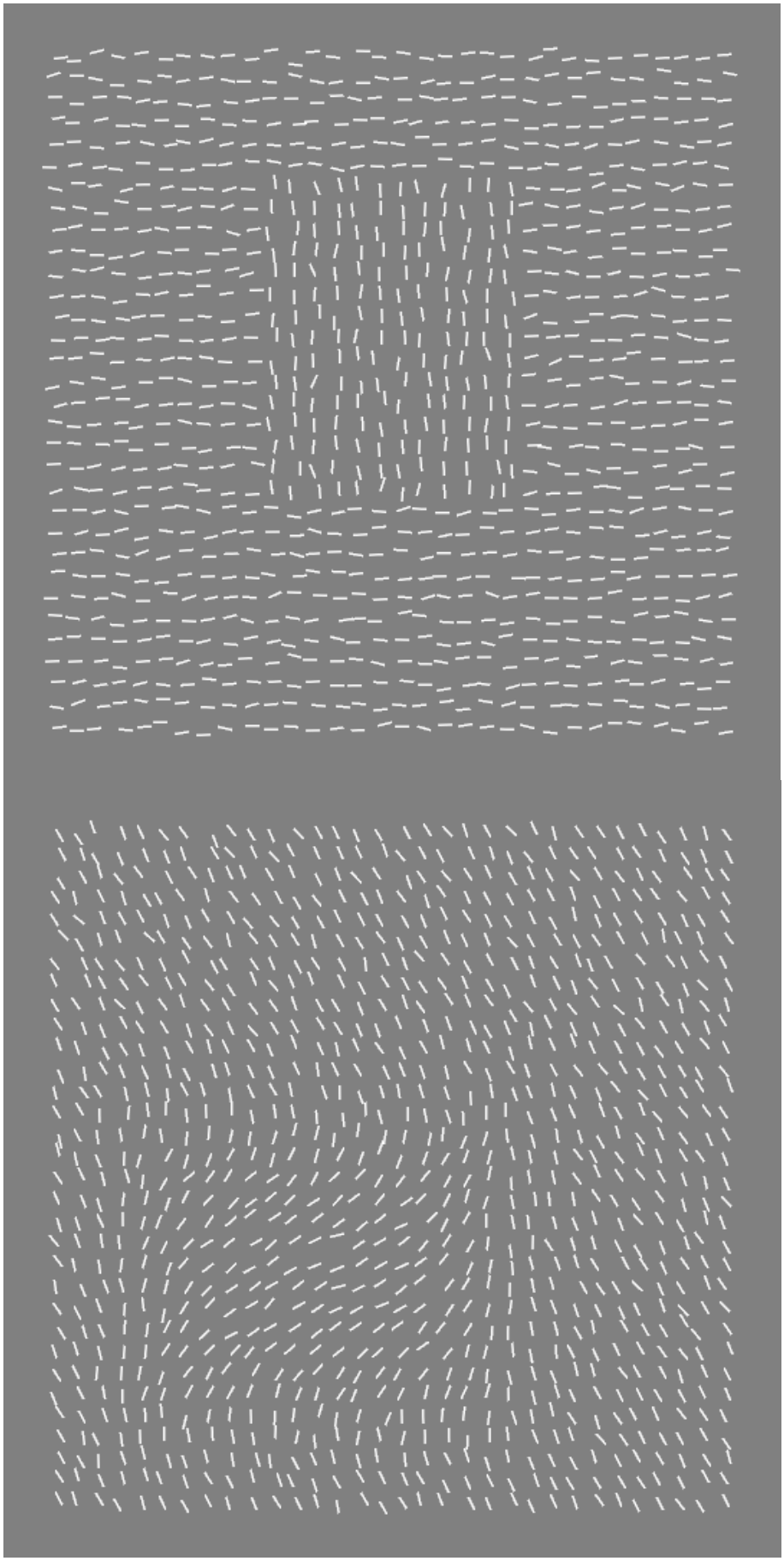
$$\theta = \frac{\Delta\theta}{1+e^{-x/k}} \quad [1]$$

$$\theta i = -\frac{\Delta\theta}{2} + \frac{\Delta\theta}{2}\sin[f(x-\frac{s}{2}) \quad [2]$$

$$\theta o = \frac{\Delta\theta}{2} + \frac{\Delta\theta}{2}\sin[f(x+\frac{s}{2}) \quad [3]$$

Where θ is the mean orientation of the line element, Δθ is the orientation contrast, x is the distance in pixels of the line element from the border (where $x = 0$), and s is the distance in pixels between the line elements. k and f are respectively the Blur and Cornsweet parameters that control the steepness of the slopes. Unless stated otherwise, k was 22.5 and f was 0.026180. Equation 1 represents the mean orientation for the Blur profile. Equations 2 and 3 represent the mean orientation for the Cornsweet profile for elements spatially positioned inside (θi) and outside (θo) the border.

The Blur profile is of a Logistic curve, in which the orientation of the line elements changes gradually over space from ground to figure. The highest degree of orientation contrast is between elements in the ground region and elements in the centre region of the figure. For the Cornsweet profile that varies spatially according to a Craik-O'Brien-Cornsweet edge (Cornsweet, 1970), the degree of orientation contrast at the immediate border of the figure is the highest, which tapers off further away from the edge. The central region of the figure and the background region have the same orientation. For every trial, the baseline orientation of the background elements was randomly chosen. Examples of the Block, Blur and Cornsweet stimuli are shown in Figure 3.3.



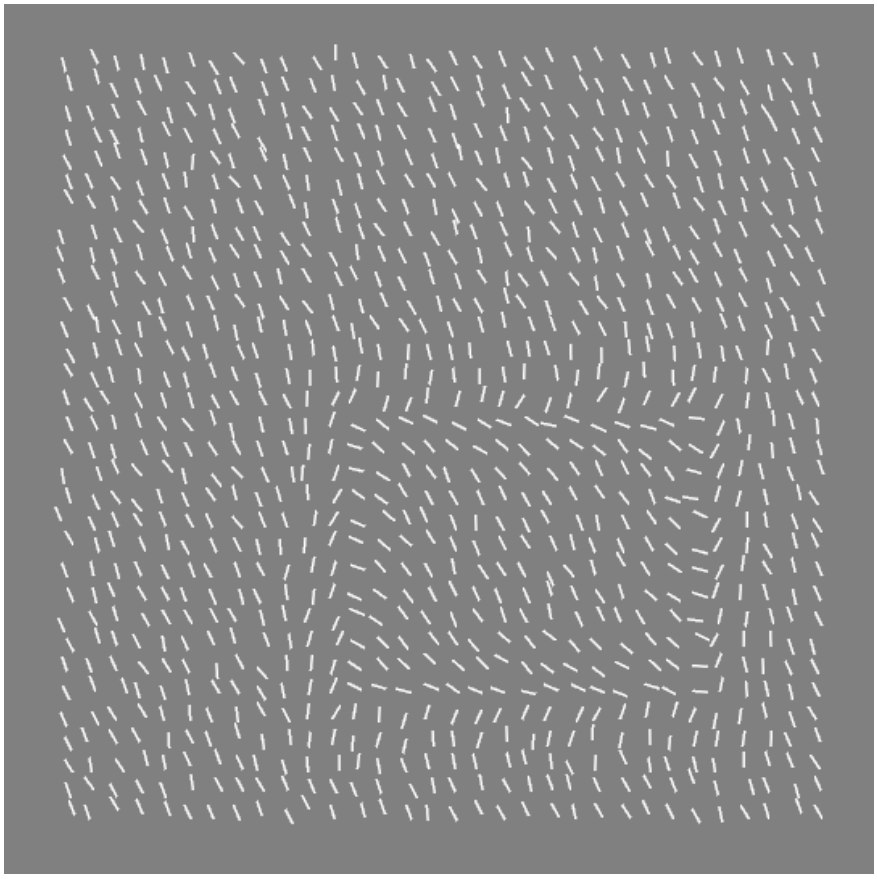


Figure 3.3. Example of the Block (top), Blur (middle), and Cornsweet (bottom) stimuli. The orientation contrast here is 90°, and the SD=7.5.

For most of the experiments, the positions of the line elements were jittered by 2 arcminutes (2 pixels) in the standard condition. The positional jitter was introduced by randomly drawing from a Gaussian distribution with a specified standard deviation (SD). The SD of the Gaussian distribution of orientation jitter was 7.5° [standard condition]. This represented the amount of external orientation noise in the stimuli (see Figure 3.4). Additionally, Figure 3.5 shows how a figure pops-out from the surround with greater ease when orientation contrast is higher.

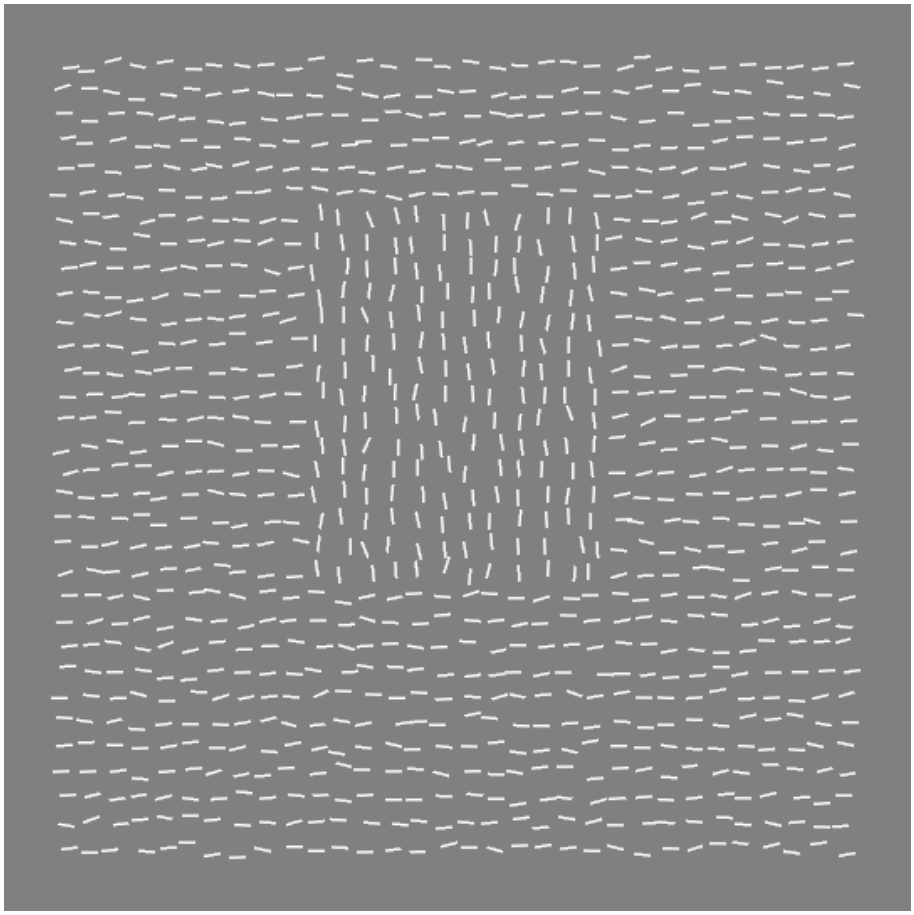


Figure 3.4. Block stimuli with orientation contrast of 90°. Orientation jitter with SD=7.5°.

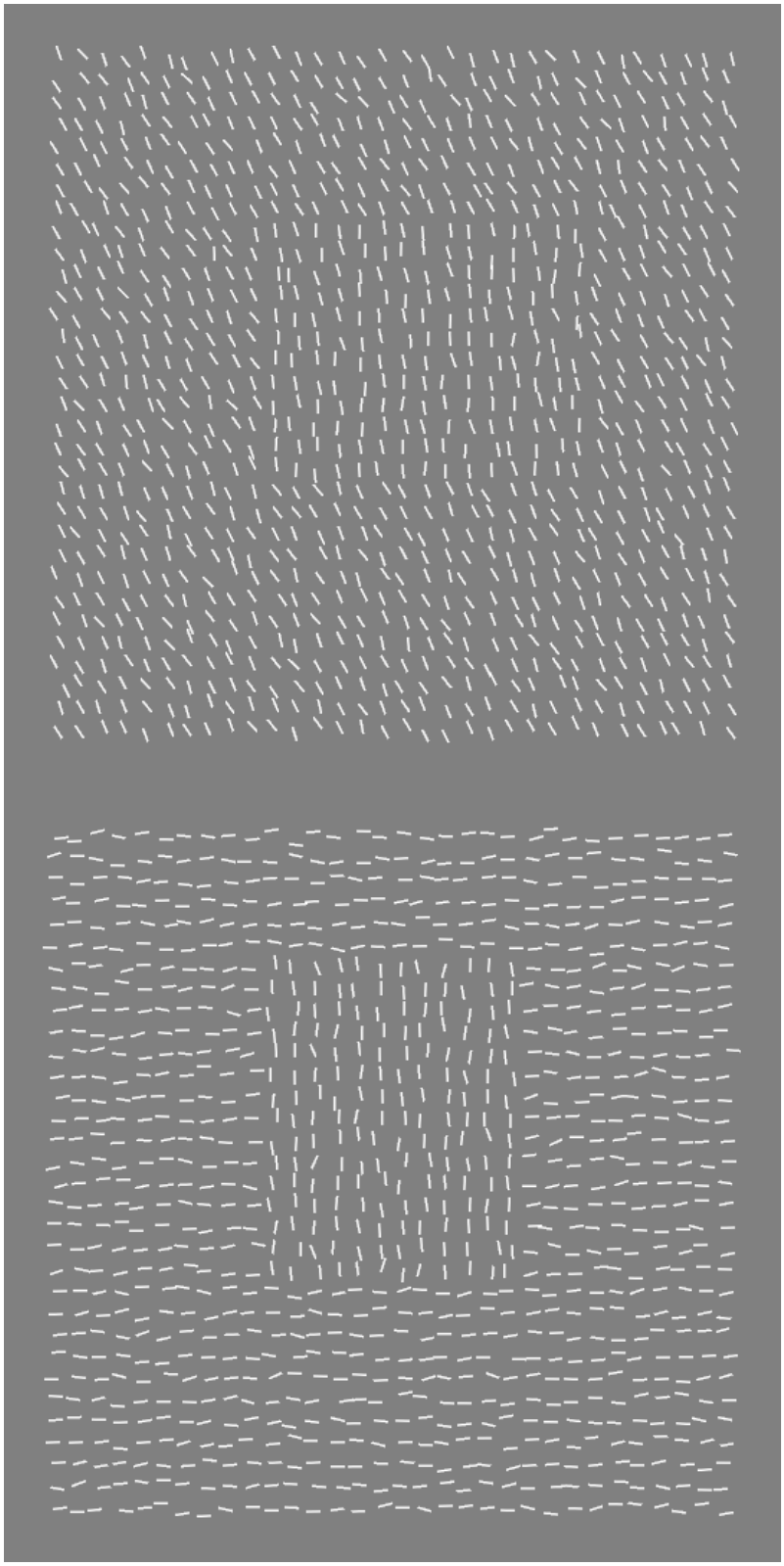


Figure 3.5. Block profile with orientation jitter of SD=7.5°, with 20° (top) and 90° (bottom) orientation contrast. Note how an increase in orientation contrast makes the figure region pop-out immediately.

For most experiments, the texture stimuli consisted of a 32×32 grid array of white line elements on a grey background. The rectangular-shaped figure could be oriented horizontally (15×12, wider) or vertically (12×15, longer). The position of this figure within the grid was varied from trial-to-trial, but with the stipulation that the border of the figure would appear after at least 5 lines from the edge of the array i.e. figure would appear on the 6th line onwards.

3.4 Tasks

For the psychophysical experiments, two different tasks were used – Segmentation (two-alternate forced-choice, 2AFC) and Detection (two-interval forced-choice, 2IFC), which will be described in more detail below. The difference between the two tasks is that the Detection task requires participants to only notice if there is a region (the figure) that looks different from the surround. For the Segmentation task, participants need to make a judgement about the orientation (vertical or horizontal) of the figure. Therefore, in the Segmentation task, it will not suffice to just notice the presence of a figure, but also to segregate the entire boundary of the figure from the background. Arguably, the Segmentation task is a true measure of segregation, as it requires the full boundary of the figure to be perceived. However, the 2IFC Detection task has also frequently been used to investigate texture segregation (e.g. Wolfson & Landy, 1998; Kingdom & Keeble, 1996; Kingdom & Keeble, 1999), and it would therefore be difficult to determine if the effect of texture segregation we have is due to the stimulus itself or the task characteristic. Thus, for most of the experiments, we used both tasks to investigate texture segregation.

Unless stated otherwise, all participants carried out several practice blocks before commencing with the experimental blocks. For the naïve observers in particular, experimental blocks began after overall accuracy reached a plateau. Observers were only

permitted to make a response after stimulus offset, and they received feedback in the form of an auditory tone for incorrect responses. The experiment was carried out in a dimly lit room.

3.4.1 Detection

Figure 3.6 shows the sequence of one trial for the detection task. The key feature of this task is that the stimulus is shown twice to the observer (2IFC), and only in one interval will there be a figure-ground texture present. In the other interval, no figure-ground texture was present i.e. average orientation of line elements is the same. The observer is required to respond with a key-press to indicate in which interval the figure-ground texture was present.

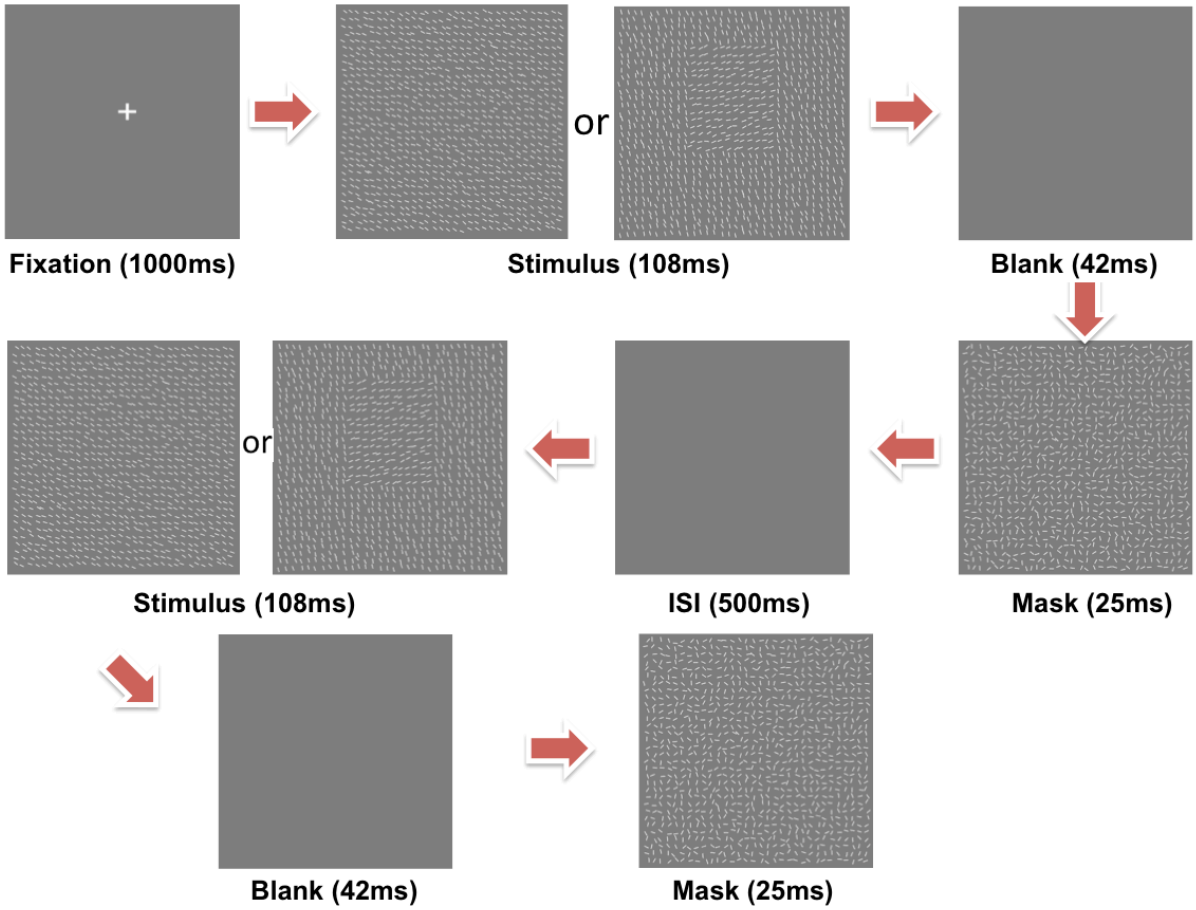


Figure 3.6. The sequence of presentations within one trial for the detection task.

3.4.2 Segmentation

For the segmentation task, observers were shown the texture stimulus once, and had to determine what the orientation (horizontal or vertical) of the figure region was. Figure 3.7 shows the sequence of one trial for the segmentation task.

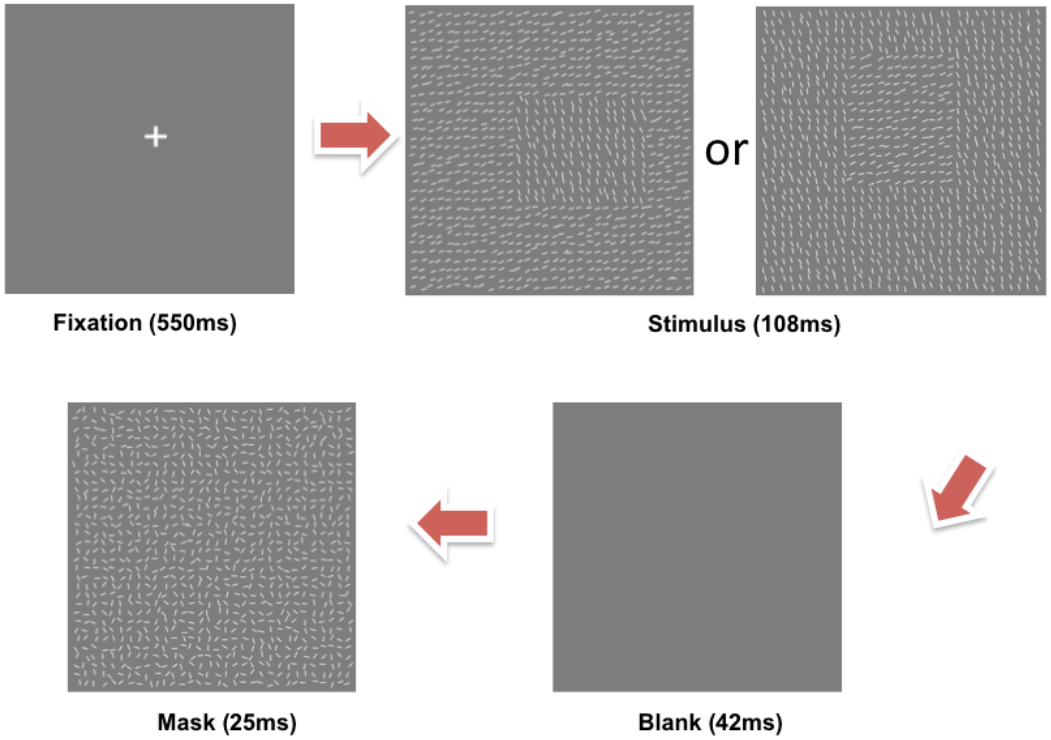


Figure 3.7. The sequence of presentations within one trial for the segmentation task.

3.5 Modeling

Besides the behavioural data we obtained from the psychophysical tasks, we also modelled the results of the studies. We did this because of the likelihood that the extraction of edge information by the visual system may extend over a larger region than the immediate border of the figure. Our model represents a mechanism that is used to average the responses over an extended region by the vicinity of the border. We will from here on refer to this as the Integration Region (IR). The rationale of calculating this IR is that a common mechanism might be used with the three different orientation profiles (Kingdom & Keeble, 1996). As

such, there would be a region size in which the orientation contrast thresholds from the three profiles will be similar to each other, and this point will be the IR.

To model the IR, we first compute the average orientation on either side of the border for the different raster units (N) using the participants' threshold results. For the Block and Cornsweet profiles, the threshold of the participant would represent the orientation contrast at the border when N=1 (i.e. orientation of line elements immediately on either side of the border). For the Blur profile, participants' thresholds are the orientation difference between background and the centre most of the figure. Therefore, the threshold for the Blur profile has to first be re-expressed as a difference in orientation at the border.

Subsequently, thresholds for all three orientation profiles are re-expressed in terms of the different mean orientations of the raster units up to N=6 (Figure 3.8). The IR is the value of N, where the thresholds for the different profiles are most similar (Figure 3.9).

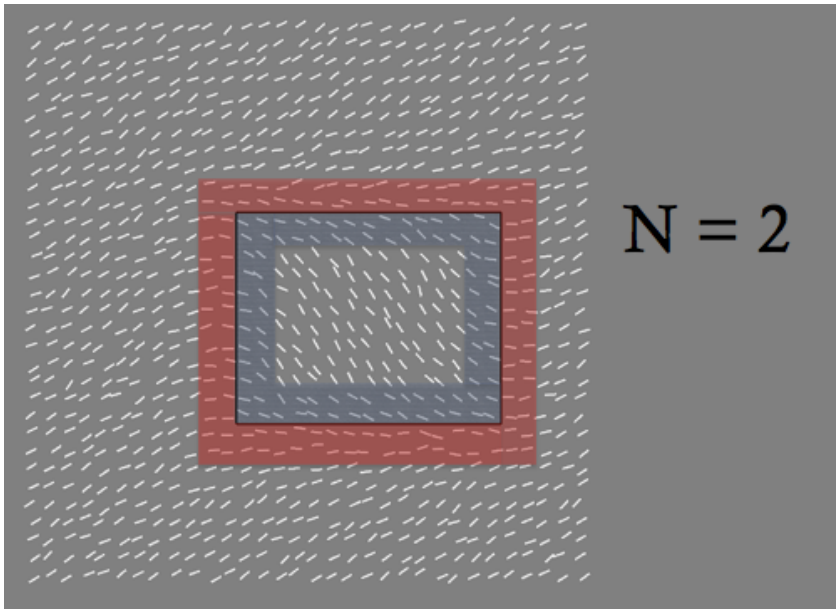


Figure 3.8. The black solid line represents the border of the figure. When N=2, the computed threshold is re-expressed in terms of an average orientation contrast on either side of the border extending 2 raster units i.e. computed threshold is an average difference between the red shaded region and the blue shaded region.

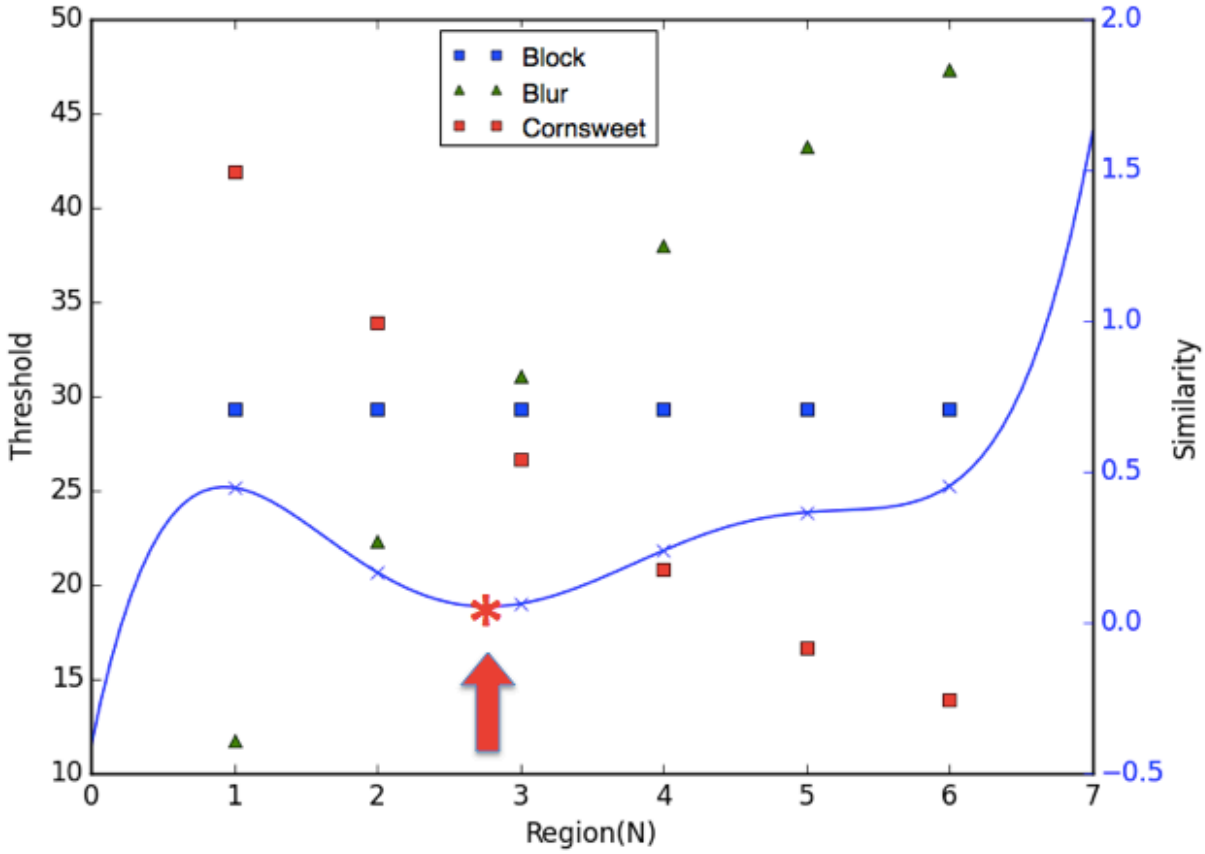


Figure 3.9. An example of how the IR was computed. The threshold results for the 3 different profiles are initially used at N=1, except for the Blur profile where the threshold is first re-expressed as the orientation contrast at the border of the figure. Subsequently, the re-expressed thresholds for the different raster widths are also computed and plotted. For each region size (N=1 – 6), the similarity between the three thresholds is calculated by dividing the standard deviation of three points (i.e. three profiles) by the mean of the three points. This is plotted as the blue x, with the scale shown of the left side y-axis. A polynomial curve is fit through these points, and the lowest point on the curve (red *, which represent the point at which the 3 orientation profile thresholds are most similar) is used as the IR. A 5th degree polynomial curve was used to fit the standard deviation, but the type of curve used here is not particularly important.

We used the bootstrap method, which is a resampling technique, to obtain error estimates of the IR size. To do this, we resampled the thresholds of all three profiles by drawing randomly from a Gaussian distribution with a mean and standard deviation that was the participant's thresholds and error estimates for the three profiles of a particular condition. We repeat these

resampling 1000 times, and the final IR and the error associated with the region size were respectively the mean and standard deviation of the 1000 samples. The error bars on IR graphs will represent the standard deviation.

3.6 Eye Tracking Methodology

For all the eye tracking studies presented in this thesis (Chapters 8 to10), the Tobii T60 eye tracker running Tobii Studio 3.2 was used to present the stimuli and record eye movement data. Prior to eye movement recording, all participants performed a 9-point calibration. The Tobii T60 samples data at 60Hz (16.67ms), with an average gaze accuracy of 0.5°, and uses the combined pupil and corneal reflection technique. The eye-tracker uses binocular tracking, with both bright and dark pupil tracking. During calibration, both methods are used, and the method that produces the highest accuracy is chosen for the experimental recording. These concepts of the eye tracker are discussed in more details at the end of Section 7.3 (Measuring Eye Movements).

3.6.1 Eye Tracking Measures

Tobii Studio was used to export the various metrics of the eye movement data to Excel for further analysis and graphing. The fixation count and summed fixation duration measure were exported directly from the built-in tools from Tobii Studio. The position of first fixation, region of first fixation, and saccade latency were calculated from the raw data outputs provided by Tobii.

Gaze point was calculated using the average position of the left and right eye. The default settings of the Tobii Studio were used to detect saccades and fixations. Saccades were classified as data points with angular velocity above 30°/s, while fixations were classified as data points with angular velocity below 30°/s. Multiple fixations, shorter than 75ms and

smaller than 0.5° visual angle, were merged into a single fixation. Fixations that do not meet the aforementioned criteria, but are also shorter than 60ms are discarded. The saccade before and the saccade after an excluded fixation were merged to form a new saccade.

3.6.1.1 First Fixation

We were able to calculate three measures from the primary fixation made by the participants. Specifically, the parameters we were interested in were the spatial "where" and temporal "when" aspects of the first fixation. For the spatial aspect, the measures we used were the position of first fixation and region of first fixation. Both of these measure the saccade landing position of the primary fixation, but the position of first fixation measure is expressed in terms of the x-coordinates along the horizontal axis, while the region of first fixation is expressed as the proportion of trials that the first fixation fell within an Area of Interest (AOI). The AOI is discussed below in section 3.6.2.

For the temporal aspect of the first fixation, we measured the saccade latency, which is the time taken to initiate a saccade after stimulus onset. It is believed that during this latency period, information is accumulated and processed in order to determine the saccade trajectory (Carpenter, 1981). Moreover, in the case of our texture stimuli, before a saccade can be made, the segregation of figure from ground has to occur first (e.g. Deubel & Frank, 1991). This measure could therefore reflect the processing required when planning a saccade to a texture stimulus.

3.6.1.2 Fixation Count

The fixation count is the measure of the number of fixations in an AOI. According to Holmqvist et al., (2011), the fixation count measure is an indication of importance (important/informative areas have a higher number of fixations), and also search difficulty

(higher number of fixations suggesting greater difficulty). Other factors such as experience, age, memory, and dysfunctions also affect the number of fixations, but these should not influence our results as we are testing healthy university students that are close in age, and have no experience observing stimuli like ours. As such, the fixation count measures will be a reflection of the various stimulus parameters.

The raw measure of the number of fixations i.e. the fixation count, were compiled for each trial per participant. These values were then averaged across all trials of the same condition per participant. The statistical analyses were carried out using these data, with the overall descriptive statistics representing the average response of all participants.

3.6.1.3 Summed Fixation Duration

The summed fixation duration measure is the total time spent fixating in an AOI (i.e. sum of all fixation durations to the AOI per stimulus presentation). Studies using face stimuli have found that the eye region is looked at for longer times compared to other parts of the face (e.g. Janik, Wellens, Goldberg, & Dell'Osso, 1978). Thus, long durations of fixations are indicative of high levels of interest to the region. Additionally, longer viewing times are also associated with greater cognitive exertion during information processing (e.g. Rayner, 1998). Additionally, an increase in time spent looking at a particular region is associated with motivation and top-down factors. For example, Rayner, Rotello, Stewart, Keir, and Duffy (2001) provided participants with print advertisements, and different groups were told to pay special attention to different types of advertisements. They found that participants spent more time looking at the specified advertisement compared to the other types of advertisements. Therefore, top-down influences, such as specific instructions or tasks, can affect the summed fixation duration.

The raw measure of the summed fixation durations were compiled for each trial per participant. These values were then averaged across all trials of the same condition per participant. The statistical analyses were carried out using these data, with the overall descriptive statistics representing the average response of all participants.

3.6.2 Data Analysis

Tobii Studio allowed for the creations of Areas of Interests (AOIs). We created three AOIs for our stimuli – centre, edge, and background (see Figure 3.10). AOIs can be useful as they are robust due to their pooled responses (i.e. all responses to the AOI are taken into account instead of single individual responses). The edge AOI is defined as the region that encompasses two columns on either side of the figure border (black line in Figure 3.10). The centre AOI is the region within the edge AOI, and the background AOI is the region beyond the edge AOI. The area of the centre AOI is 19800 pixels^2 (10.9°^{2} for experiments in Chapter 8, and 12.6°^{2} for experiments in Chapter 9), while the area of the edge AOI is 48600 pixels^2 (26.8°^{2} for experiments in Chapter 8, and 30.9°^{2} for experiments in Chapter 9). For the experiments presented in Chapter 8, the texture stimuli were of 32×32 grids, which had a ground area of 162000 pixels^2 (89.2°^{2}). For the experiments presented in Chapter 9, a 50×68 texture grid was used, with the ground AOI having an area of 696600 pixels^2 (443.3°^{2}). The size of the ground area does not include the centre and edge area.

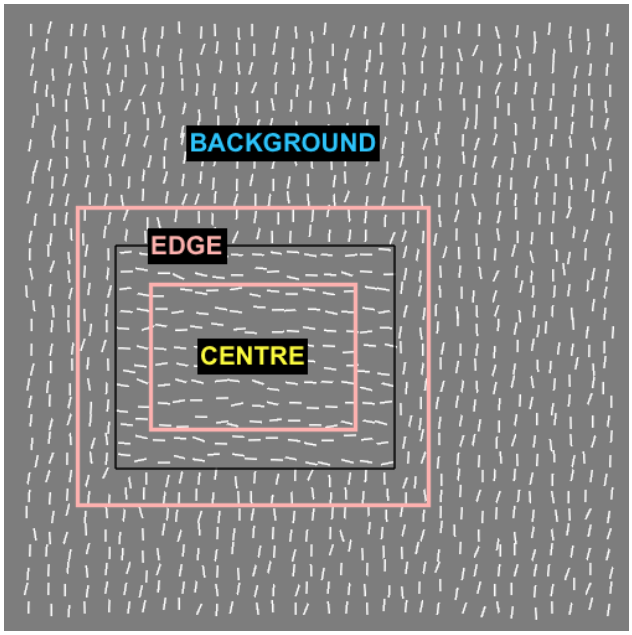


Figure 3.10. Example of the pre-defined regions of interest for the 32 by 32 texture grid. The solid black line represents the border of the figure.

Considering how the chance of fixating on a particular AOI is dependent not only on the content of the stimulus, but also the size of the stimulus (e.g. Holmqvist et al., 2011; Tatler, 2007), we needed to account for the different sizes of our AOIs. Therefore, for measures that were analysed in terms of proportion response to the different AOIs, each raw measure was transformed into area-normalized responses (e.g. Fletcher-Watson et al., 2008; Bindemann et al., 2009).

This was achieved by initially calculating proportion responses for the different measures, which were expressed as the percentage of response to each AOI. Subsequently, those proportions were divided by the size of the AOI, which was expressed as a percentage of the total stimulus size. The rationale behind this is as follows. If the fixation proportion of a region is 50%, and this region occupies 50% of the screen, the area-normalized score will be one (1), which indicates that the fixations to that region are occurring at the same rate as would occur by chance. However, scores greater than 1 (e.g. 50% fixation time to a region

which occupies 10% of the screen that results in a score of 5) would indicate fixations were directed to the region more than randomly predicted. Importantly, this normalization allows us to make direct comparisons between the different AOIs, without the influence of AOI size. Therefore, area-normalized responses were only used for analysis into the different AOIs.

Statistical analyses were conducted using SPSS. Unless stated otherwise, data were analysed using repeated measures ANOVAs. For all statistical tests conducted, when the assumption of sphericity had been violated, the Greenhouse-Geisser or Huynd-Feldt correction was applied to the degrees of freedom and the p-value. The Greenhouse-Geisser correction was used when ε was less than 0.75, and the Huynd-Feldt correction when ε was more than 0.75. Furthermore, only when main effects and interactions were significant, post hoc comparisons were conducted. Bonferroni corrections (adjusted p-values) were applied for all post hoc tests.

Chapter 4

The Role of Texture Edges in Segmentation and Detection Tasks

From the literature reviewed in Chapter 2, it can be said that our understanding of the mechanisms used for texture segregation is still incomplete. There are several competing ideas (edge-based or region-based mechanisms) in the literature, with no conclusive evidence for one or the other. Our study aims to address the question of the different mechanisms used during texture segregation, and also explore the role of texture edges in segregation. This was studied in Experiment 1, with behavioural results and modeling data suggesting that texture segregation occurs due to a fixed-size second-order filter i.e. no indication of a filling-in mechanism. In Experiment 2, we tested the possibility of texture segregation occurring due to a fixed-size filtering mechanism, and the results were in line with what one would expect should texture segregation be due to the aforementioned mechanism.

4.1 Experiment 1: Effect of Orientation Contrast at the Edge and Centre as a Function of Time

In this experiment, participants performed both the detection and segmentation task, with the three different types of stimuli – Block, Cornsweet and Blur. The key characteristic of these types of stimuli is the different information they present on the edge and centre of the figure. The Block profile has information of orientation contrast at both the edge region and centre region of the figure. The Cornsweet profile has orientation contrast information only at the edge region, with the orientation of the line elements at the figure centre and background being the same. For the Blur profile, there is a gradual change in orientation of the line elements from the background to the central region of the figure, with maximal orientation

contrast occurring between figure centre and ground. For the Blur profile, one would expect that region classification would more likely occur than boundary/edge detection, as there is minimal information of orientation contrast at the edge.

If texture segregation occurs due to an edge-based mechanism, it would be expected that the thresholds for the Block and Cornsweet profile will be similar, as both profiles have information of orientation contrast at the edge of the figure. Additionally, if edge extraction precedes filling-in, it would be expected that performance for the Block and Cornsweet profiles will be similar at short stimulus durations when filling-in would not yet have occurred. However, at longer stimulus durations, performance for the Block profile would be better compared to the Cornsweet profile owing to the information of orientation contrast present at the centre of the figure. On the other hand, if performance for the Block profile is always better compared to the Cornsweet profile regardless of stimulus duration, it might imply that a mechanism operating outside of the edge boundary plays a role in the segregation process. If texture segregation were due to a region-based mechanism, we would expect that the thresholds for the Block and Blur profile would be similar to each other, as both have information of orientation contrast at the centre of the figure, while performance for the Cornsweet profile will be poorer.

4.1.1 Methods

4.1.1.1 Participants

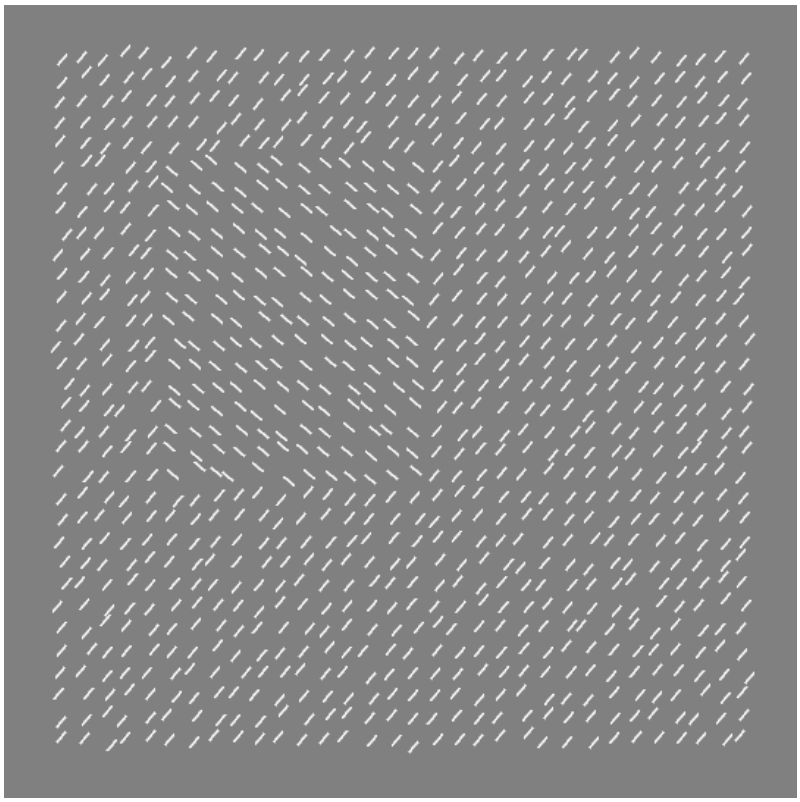
Three observers (DK, SKS and HK) participated in the experiment. DK and SKS were experienced psychophysical observers, while HK was a naive observer. All participants had corrected-to-normal vision.

4.1.1.2 Display

The experiment was coded using C on a Macintosh Quadra 840AV computer, and displayed on a gamma-corrected CTX CRT monitor. The monitor had a refresh rate of 74.5Hz. The algorithms used to produce the anti-aliased lines were in accordance with Gupta and Sproull (1981). At the viewing distance of 121 cm, the pixel resolution was 1 pixel per arcminute.

4.1.1.3 Stimuli

The stimuli were created as described in Chapter 3, Section 3.3. However, in this experiment, orientation jitter was varied, whereby the standard deviation (SD) could be either 0°, 7.5° [standard condition, as described in Section 3.3], or 15° (see Figure 4.1). The entire texture grid spanned 8.0°×8.0°, while the 12×15 figure spanned 3.0°×3.8°.



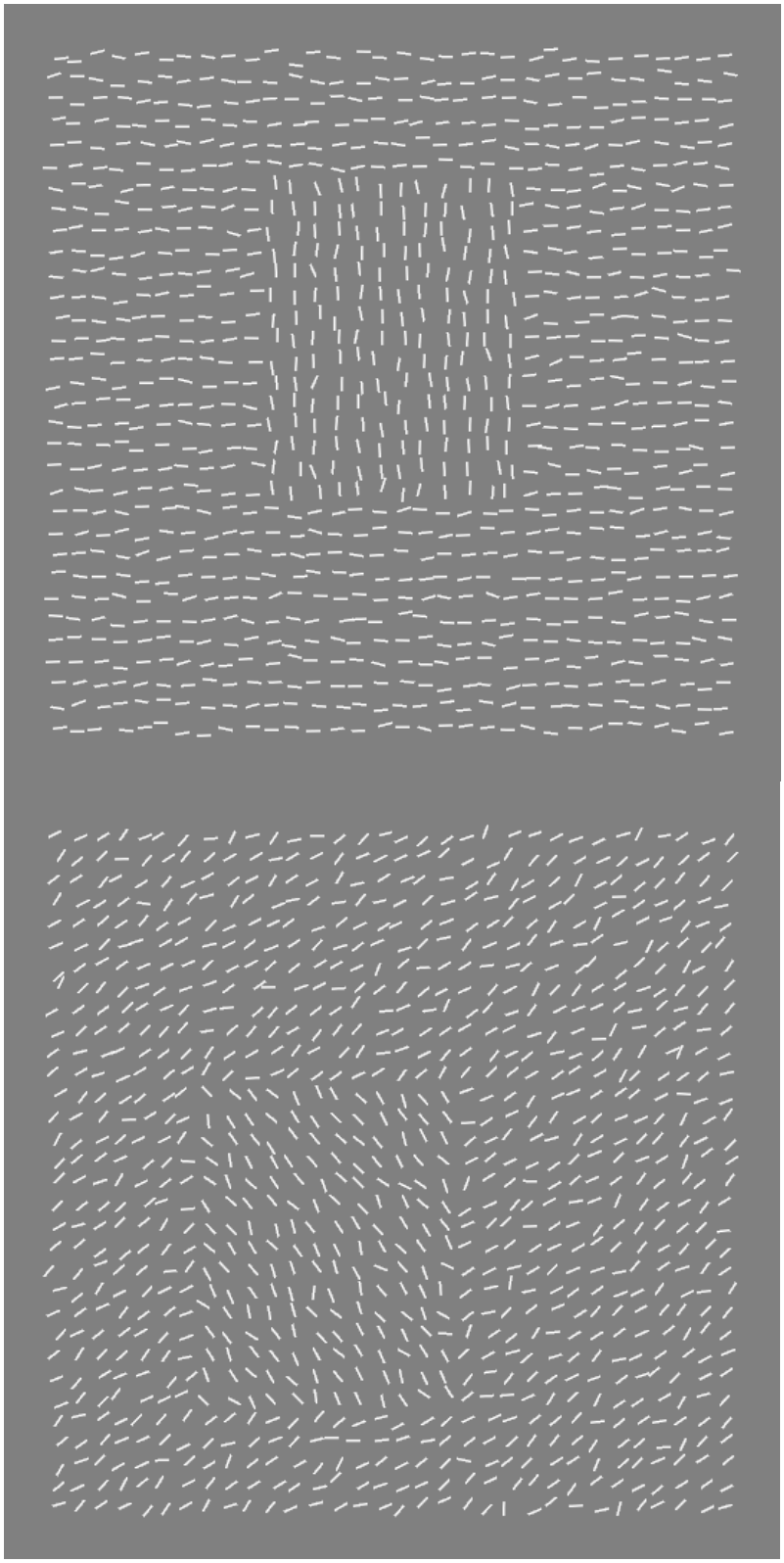


Figure 4.1. Block stimuli with orientation contrast of 90°. Orientation jitter with SD=0° (top), SD=7.5°(middle), and SD=15° (bottom). Increasing the orientation jitter makes the segregation of figure from ground more difficult.

4.1.1.4 Procedure

Participants performed two tasks – Detection and Segmentation in this experiment (see Section 3.4). Briefly, the detection task is a 2IFC paradigm where two displays were presented in one trial, and participants had to determine in which interval was the figure-ground texture present. The segmentation task is a 2AFC paradigm where there was only one presentation of stimulus, and participants had to determine the orientation – horizontal/ vertical – of the figure. In this study, the ISI for the detection task was 1500ms. Additionally, for both the detection and segmentation task, five stimulus durations were used: 14, 27, 54, 108 [standard condition], and 216ms, and three orientation jitters: SD = 0°, 7.5° [standard condition], or 15°.

The method of constant stimuli was used in this study (see Section 3.1.1). During a given block of 180 trials, only one task and one stimulus duration were presented. Within each block, the different orientation profiles and orientation jitter were presented in randomised order. Each block was repeated 10 times, with a total of 100 blocks run per participant over multiple days.

4.1.2 Results

For this study, two threshold values were generated for the Blur profiles – Blur-centre and Blur-edge. For the Blur-centre, the threshold was of orientation contrast between the background and central region of the figure where orientation contrast with respect to the background was the largest. Blur-edge on the other hand was of orientation contrast between the line elements immediately adjacent to the figure border (red line in Figure 3.2, Chapter 3). Essentially, we represented the threshold of the Blur profile as a function of orientation contrast at the edge (Blur-edge) **and** orientation contrast between centre and background

(Blur-centre). The purpose of representing the Blur threshold as a function of edge orientation contrast was to ascertain whether or not an edge-based mechanism was being used. That is, should texture segregation be mediated by an edge-based mechanism, thresholds for the Blur profile as a function of edge orientation contrast (Blur-edge) should be similar to the Block and Cornsweet thresholds.

The threshold for the Blur-edge was calculated based on the threshold of the Blur-centre, using the formula:

$$\Delta\theta_e = \Delta\theta_g\left(\frac{1}{1 + e^{s/2k}} - \frac{1}{1 + e^{-s/2k}}\right)$$

Where $\Delta\theta_e$ and $\Delta\theta_g$ are respectively the threshold of orientation contrast for the Blur profile as a function of the edge and the centre of the figure. *s* is the distance in pixels between the line elements, and *k* is the Blur parameter that controls the steepness of the slope.

For the Block and Cornsweet profiles, threshold values were of orientation contrast at the figure edge border.

4.1.2.1 Effect of Stimulus Duration

A trend in the results (see Figures 4.2 – 4.4 for individual participants' data and Figure 4.5 for mean data across participants) is that orientation thresholds are lower with increased time to view the stimulus. This is found for all participants regardless of the different tasks, orientation profiles, or orientation jitter used. That is, on average, thresholds to the shortest stimulus duration are 2.9× higher than the longest stimulus duration. The data also finds that at 14ms stimulus duration, participants were mostly unable to perform either of the tasks. Only one participant was able to produce physically meaningful thresholds (i.e. <90°, as

orientation contrast greater than 90° is not possible) for the Detection task at 14ms stimulus duration.

Furthermore, for all stimulus durations, the Block profile had lower thresholds compared to the Cornsweet profile. On average, this increase in threshold for the Cornsweet stimuli was between 23% and 39% for all 5 stimulus durations tested. This result does not support the notion of an edge-based mechanism operating solely, as both the Block and Cornsweet profiles have the same information of orientation contrast at the edge, yet performance for the Block profiles is better than the Cornsweet profile.

4.1.2.2 Effect of Orientation Profiles

As can be seen from the data plots of all 3 participants (Figures 4.2 – 4.4, and Figure 4.5 for group mean), performance for the Blur profile as a function of edge contrast was the best (*M*=9.9°), followed by the Block profile (*M*=38.9°), Cornsweet profile (*M*=46.4°), and finally the Blur profile as a function of centre contrast (*M*=59.9°). Further observations of the plots for all 3 participants show that for the detection task, only participant HK exhibited a 23% decrease in threshold values for the Cornsweet profile compared to the Blur-centre thresholds, while the other participants had similar threshold values for the Cornsweet and Blur-centre (decrease in threshold by 2.6% for the Cornsweet profile). However, for the segmentation task, orientation thresholds for the Blur profile as a function of centre contrast was on average 45.7% higher than the Cornsweet profile for all participants. This will be discussed further in [Effect of Task].

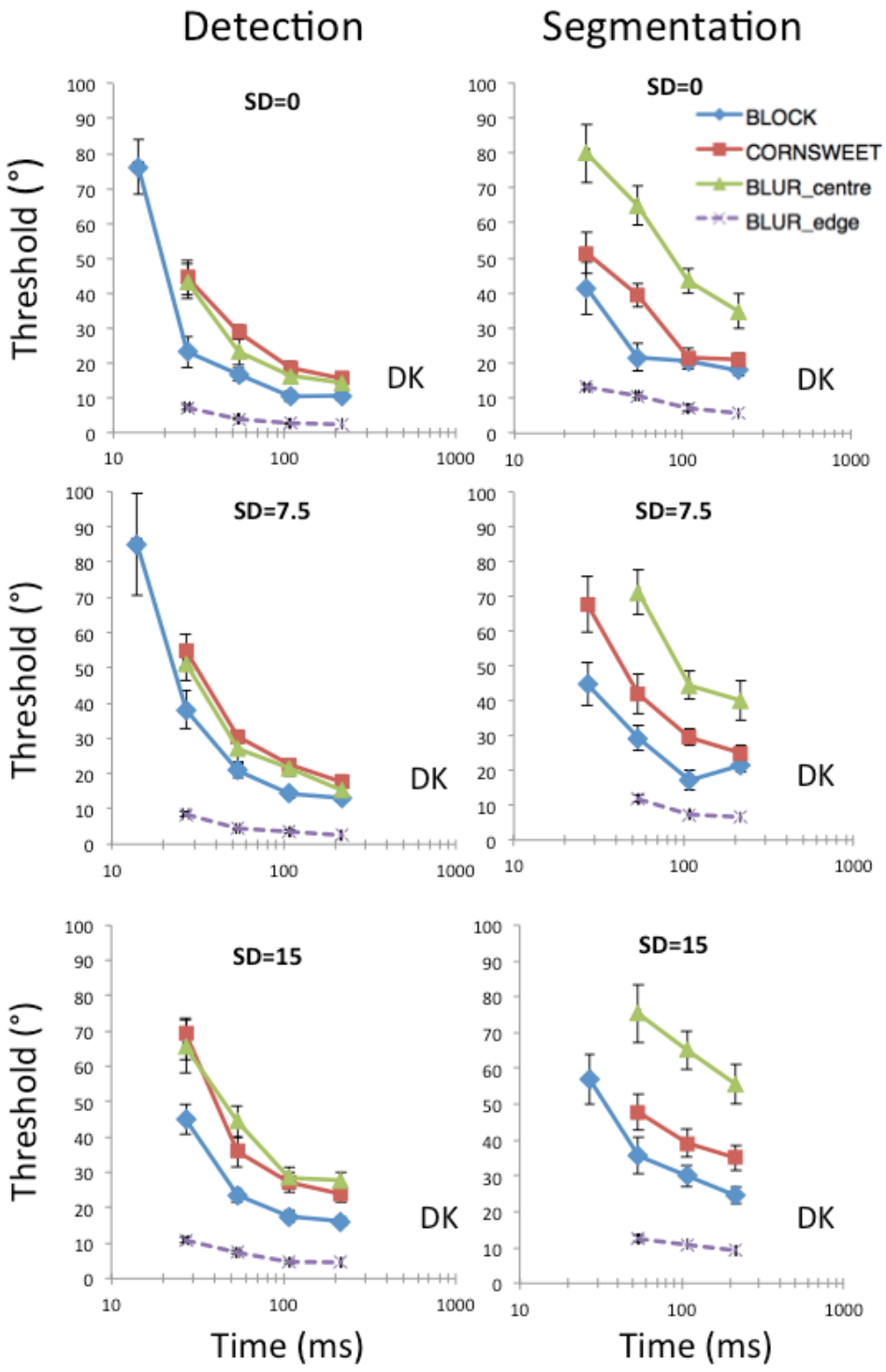


Figure 4.2. Thresholds values for participant DK as a function of stimulus duration for the 3 orientation profiles. Thresholds for the Blur profile defined by orientation contrast at the edge are plotted in dashed lines. Differing orientation jitter is plotted on the different rows for both the detection (left column) and segmentation (right column) tasks. Error bars represent the 67% confidence interval.

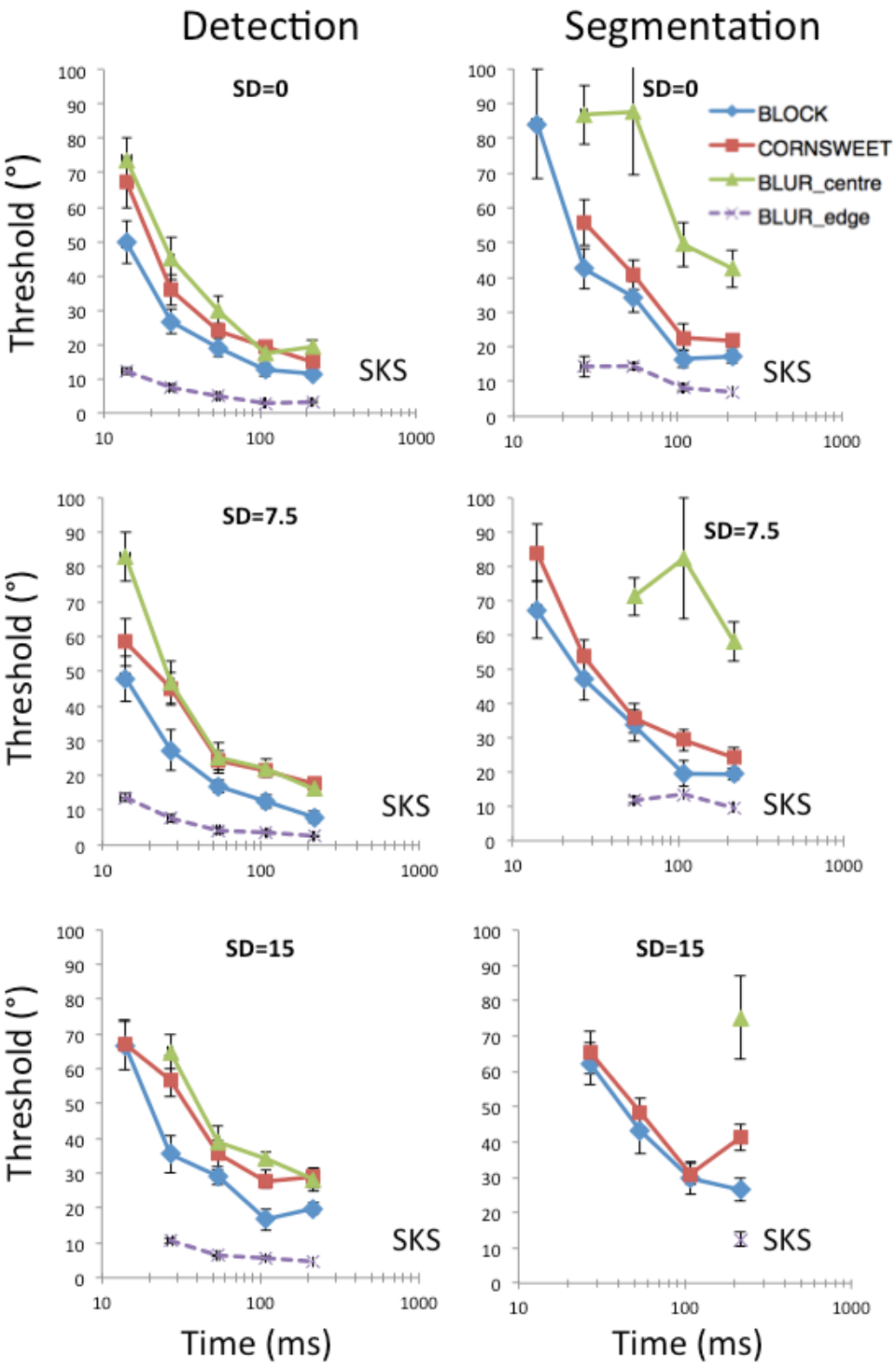


Figure 4.3. Thresholds values for participant SKS as a function of stimulus duration for the 3 orientation profiles. Thresholds for the Blur profile defined by orientation contrast at the edge are plotted in dashed lines. Differing orientation jitter is plotted on the different rows for both the detection (left column) and segmentation (right column) tasks. Error bars represent the 67% confidence interval.

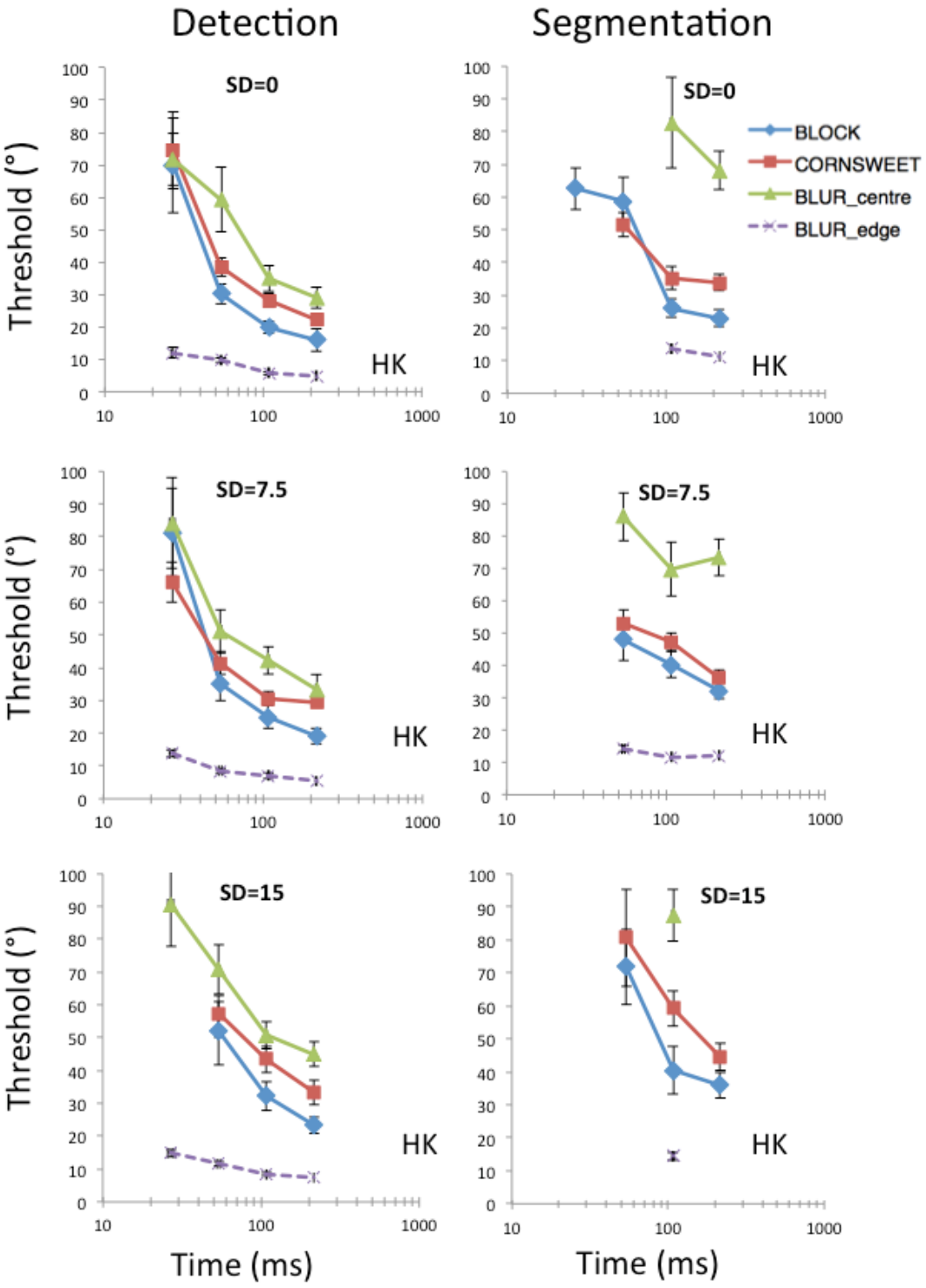


Figure 4.4. Thresholds values for participant HK as a function of stimulus duration for the 3 orientation profiles. Thresholds for the Blur profile defined by orientation contrast at the edge are plotted in dashed lines. Differing orientation jitter is plotted on the different rows for both the detection (left column) and segmentation (right column) tasks. Error bars represent the 67% confidence interval.

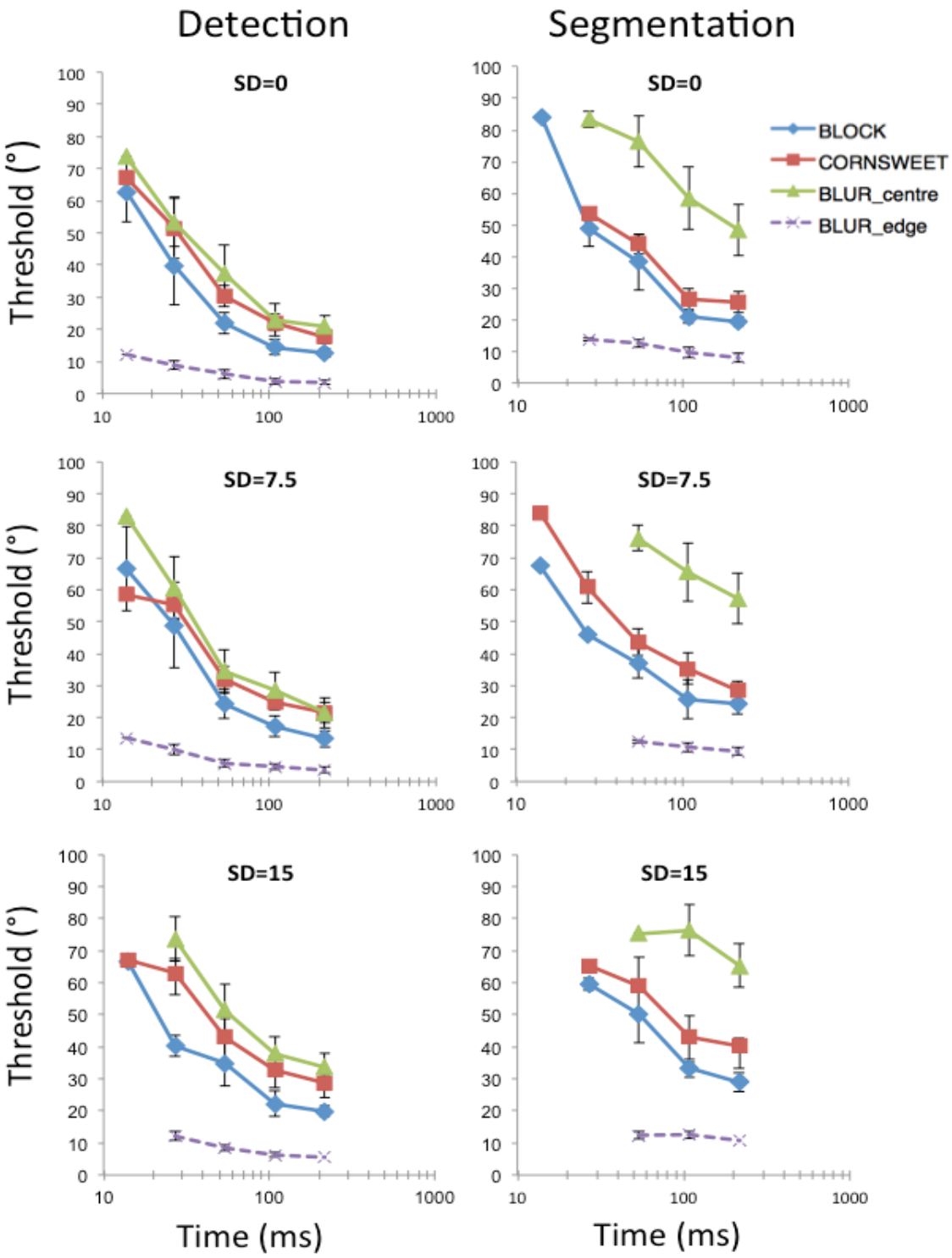


Figure 4.5. Average thresholds for all participants (*n*=3) as a function of stimulus duration for the 3 orientation profiles. Thresholds for the Blur profile defined by orientation contrast at the edge are plotted in dashed lines. Differing orientation jitter is plotted on the different rows for both the detection (left column) and segmentation (right column) tasks. Error bars represent the standard error of the mean.

4.1.2.3 Effect of Orientation Jitter

Performance was worse (i.e. increased orientation threshold) with increased orientation jitter (note the linear increase in the data plots of Figure 4.6). However, the orientation jitter has not influenced the outcomes of the other results. That is, with all orientation jitters tested, the effects of stimulus duration (better performance with increased time) and orientation profiles (best performance for the Block profile) are maintained. Therefore, we can be certain that the conclusions we have drawn about the results are not critically dependent on the particular standard deviation used. Additionally, this effect of orientation jitter on performance could be explained in terms of the signal detection theory (Macmillan & Creelman, 2005). The increased orientation jitter would be an increased source of external noise, which contributes to the additional uncertainty in making a response or judgement about the stimulus.

4.1.2.4 Effect of Task

Overall, orientation thresholds for the detection task were lower by 32% compared to the segmentation task. Apart from that, the different tasks have an effect on the orientation thresholds of the different orientation profiles. For both tasks, the Blur profile as a function of edge contrast had the lowest orientation thresholds, followed by the Block profile. However, orientation thresholds for the Cornsweet and Blur-centre were similar for the detection task (for 2 out of 3 participants, there was only a difference of 2.6%), while orientation thresholds for the Blur-centre were on average 45.7% higher than the Cornsweet thresholds for the segmentation task (see Figures 4.2 – 4.5). Thus, only for the segmentation task is the performance of the Blur profile poorer compared to the Block and Cornsweet profiles. This may well be due to the segmentation task requiring an explicit border of the figure to be formed first before a response can be made, and unlike the Block and Cornsweet profiles, the Blur profile does not have an abrupt texture edge to aid in localising the figure border.

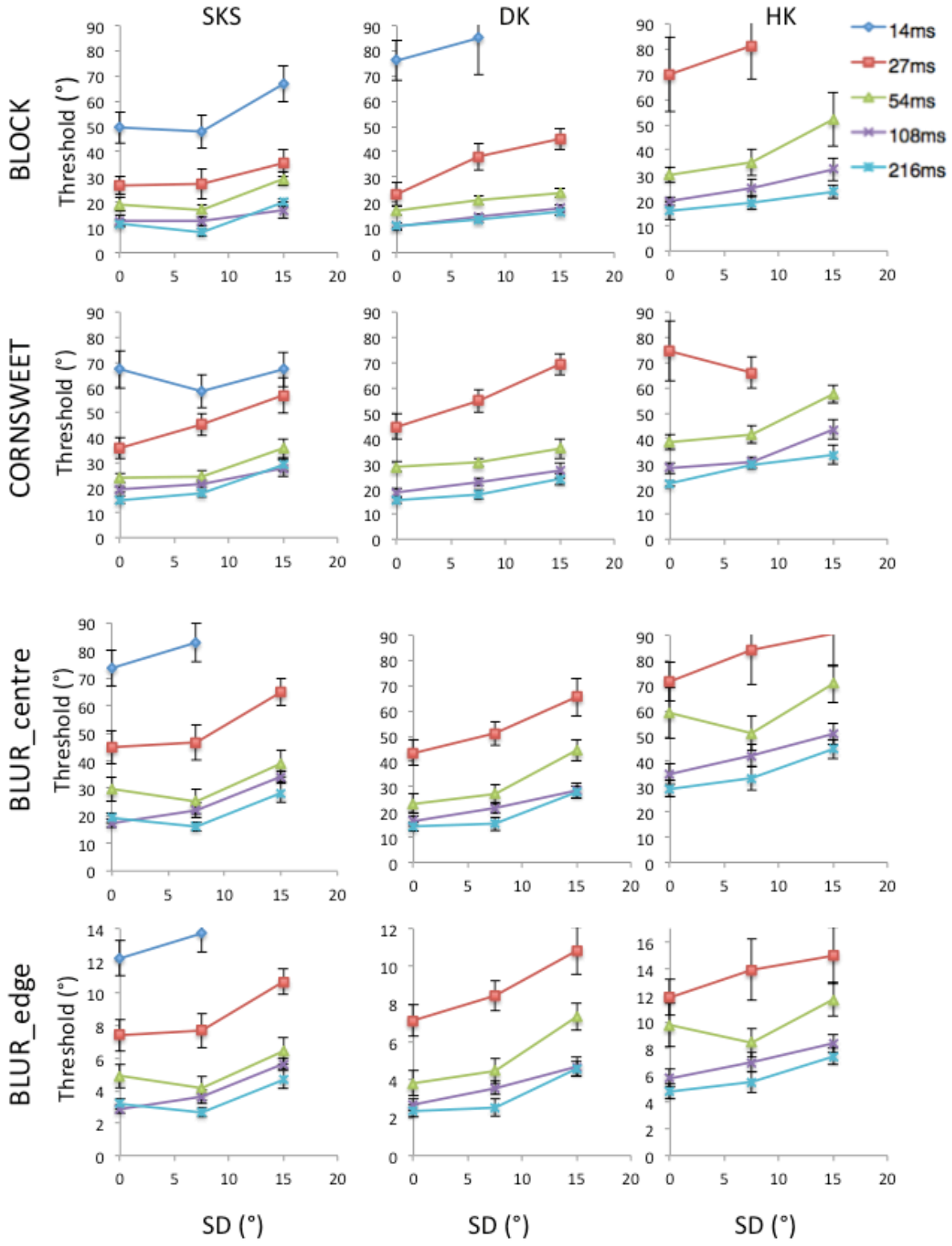


Figure 4.6. The effect of orientation jitter on orientation threshold for the Block, Cornsweet, and Blur (as a function of both edge and centre contrast) profiles for all 3 participants for the detection task. Error bars represent the 67% confidence interval.

4.1.3 Modeling

In the experiment presented, the role texture edges play in the segmentation and detection process was investigated. The study measured performance of three participants using three different profiles – Block, Blur, and Cornsweet – while performing a segmentation and detection task. The effects of stimulus duration (14, 27, 54, 108, and 216ms) and orientation jitter (SD=0°, 7.5°, and 15°) on performance were also investigated.

The results of the study show that: 1) better performance was found for the Block profile, followed by the Cornsweet profile, and worst performance for the Blur profile (Blur-centre), 2) performance improved with increased stimulus duration, 3) performance declined with increased orientation jitter, and 4) performance was better for the detection task compared to the segmentation task.

The key finding of the study was the effect of orientation profiles and stimulus duration on performance. Studies (e.g. Caelli, 1985; Caputo, 1998; Lamme, Rodriguez-Rodriguez, & Spekreijse, 1999) have suggested that segregation occurs due to a filling-in mechanism. Should this be true, thresholds for the Block profile should be better than the Cornsweet profile at longer stimulus durations, owing to the additional orientation contrast information present in the centre of the figure for the Block profile. At short stimulus durations on the other hand, the Block and Cornsweet profiles should have similar thresholds as filling-in would not yet have occurred, and both profiles have the same orientation contrast information at the edge. However, we do not find this pattern of results, which suggest that filling-in did not occur.

The results of the experiment, which show that performance for the Block profile was the best, followed by the Cornsweet, then Blur profile, imply that information about orientation

contrast is not retrieved solely by either an edge or region-based mechanism. The results also suggest that a filling-in mechanism is not aiding in texture segregation. We propose that the extraction of edge information by the visual system may well extend over a larger region than the immediate border of the figure, and we refer to this as the integration region (IR). Briefly, we calculate this IR on the assumption that a common mechanism is used to process the three different orientation profiles (Kingdom & Keeble, 1996). As a result of this mechanism, there would be a region size in which the orientation contrast thresholds from the three profiles will be similar to each other, and this point will be the IR. Using this computational model, we assume that the IR size is reflective of the size of the second-order texture filter that mediates texture segregation. For more details about this model, see Section 3.5.

For Experiment 1, we modelled the IR for both tasks and orientation jitters as a function of stimulus duration (see Figure 4.7). The IR is calculated for each condition separately (2 tasks – segmentation and detection; 3 SD – 0°, 7.5°, and 15°; 5 stimulus durations – 14, 28, 56, 108, and 216ms). For some conditions, IRs could not be calculated because thresholds were not available for all three profiles (i.e. if thresholds were above 90°). The modeling results show that orientation jitter and stimulus duration do not have much of an effect on the IR. However, the IR is different between the two tasks, with the size of the IR being higher for the detection task (~3.6) compared to the segmentation task (~2.6). The result, which shows that the size of the IR does not increase with increased stimulus duration, suggests that filling-in is not occurring. In fact, it suggests that texture segregation is occurring due to a second-order filtering mechanism (i.e. second stage filtering from the Filter-Rectify-Filter model) that has a fixed size. This is discussed further in Experiment 2 and the discussion.

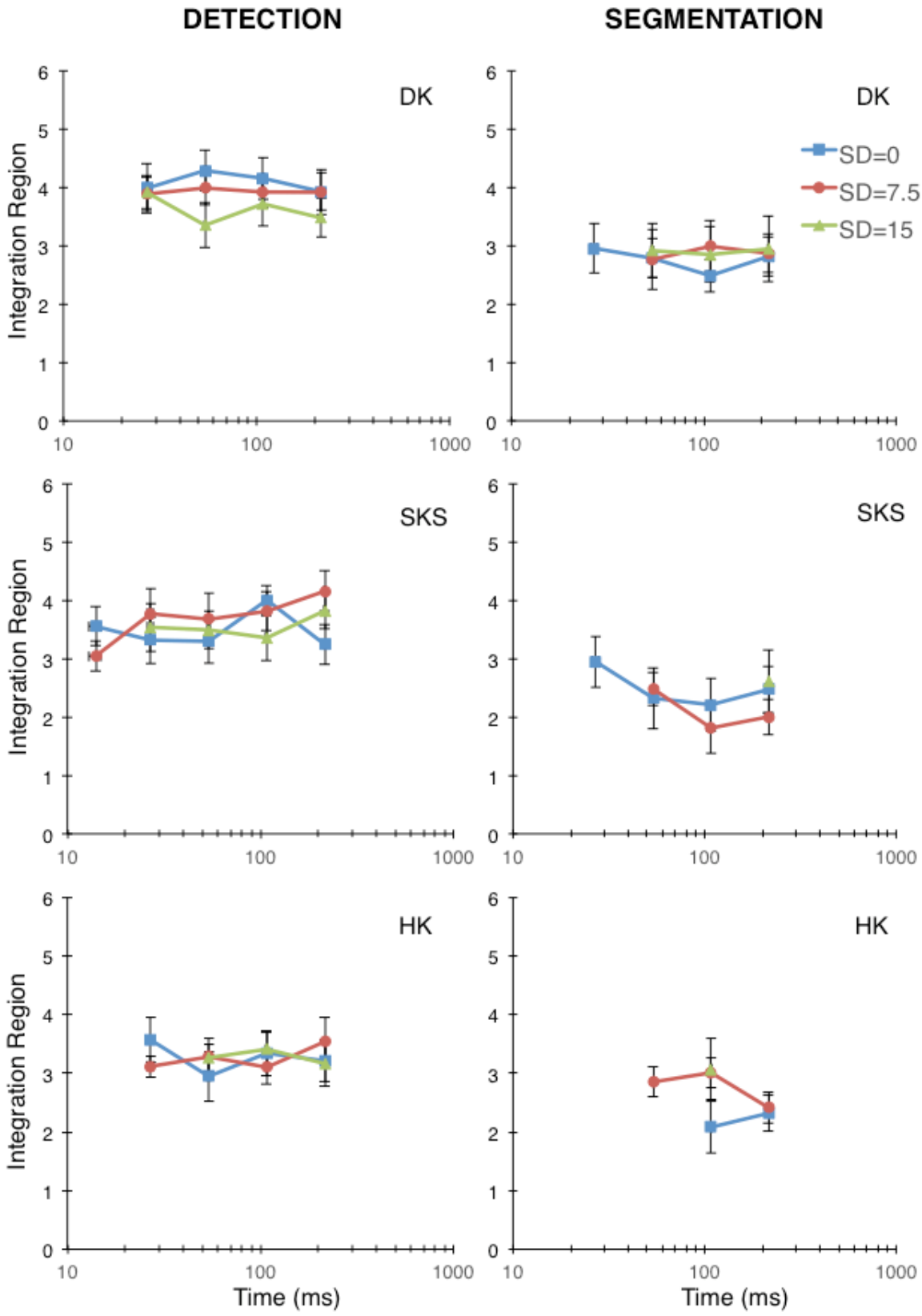


Figure 4.7 IRs for the participants for the various tasks and conditions. For conditions in which thresholds were not available for all three profiles, IR could not be computed. Error bars are the standard deviation, which were obtained by bootstrapping the data.

4.1.4 Re-creating Experiment 1

Experiment 1 was coded using C on a Macintosh Quadra 840AV computer. However, it was not feasible to continue using this setup as the Macintosh Quadra 840AV was an old piece of hardware, and we ran the risk of the computer malfunctioning/breaking down completely if we ran further experiments on that setup. Besides that, Experiment 1 was coded in the Quadra 840AV system itself, and it was not possible to extract those codes to get it operating on a new system. In the interest of furthering that line of research, we therefore had to re-create the experiment using an entirely new hardware (Apple Mac Mini) and software (PsychoPy (Pierce, 2007)).

We chose PsychoPy as we could use the built-in functions (e.g. trial randomisation, collecting responses), while **also** using the coding component (Python language) to create the complex stimuli needed for our experiments. As we had to re-script the entire experiment in a different programming language, we needed to make certain that the new experiment created was truly the same as Experiment 1.

SKS, who participated in Experiment 1, served as control-check to ensure that the results in the re-created experiment matched the results of Experiment 1. SKS performed both the detection and segmentation task at 108ms stimulus duration. All three orientation profiles were used, with the orientation jitter of SD=7.5°. Orientation contrast thresholds between Experiment 1 and the re-created experiment were compared (see Figure 4.8). Based on the similar thresholds obtained, and the overlapping error bars, we are certain that we have been successful in reproducing Experiment 1 in the new setup.

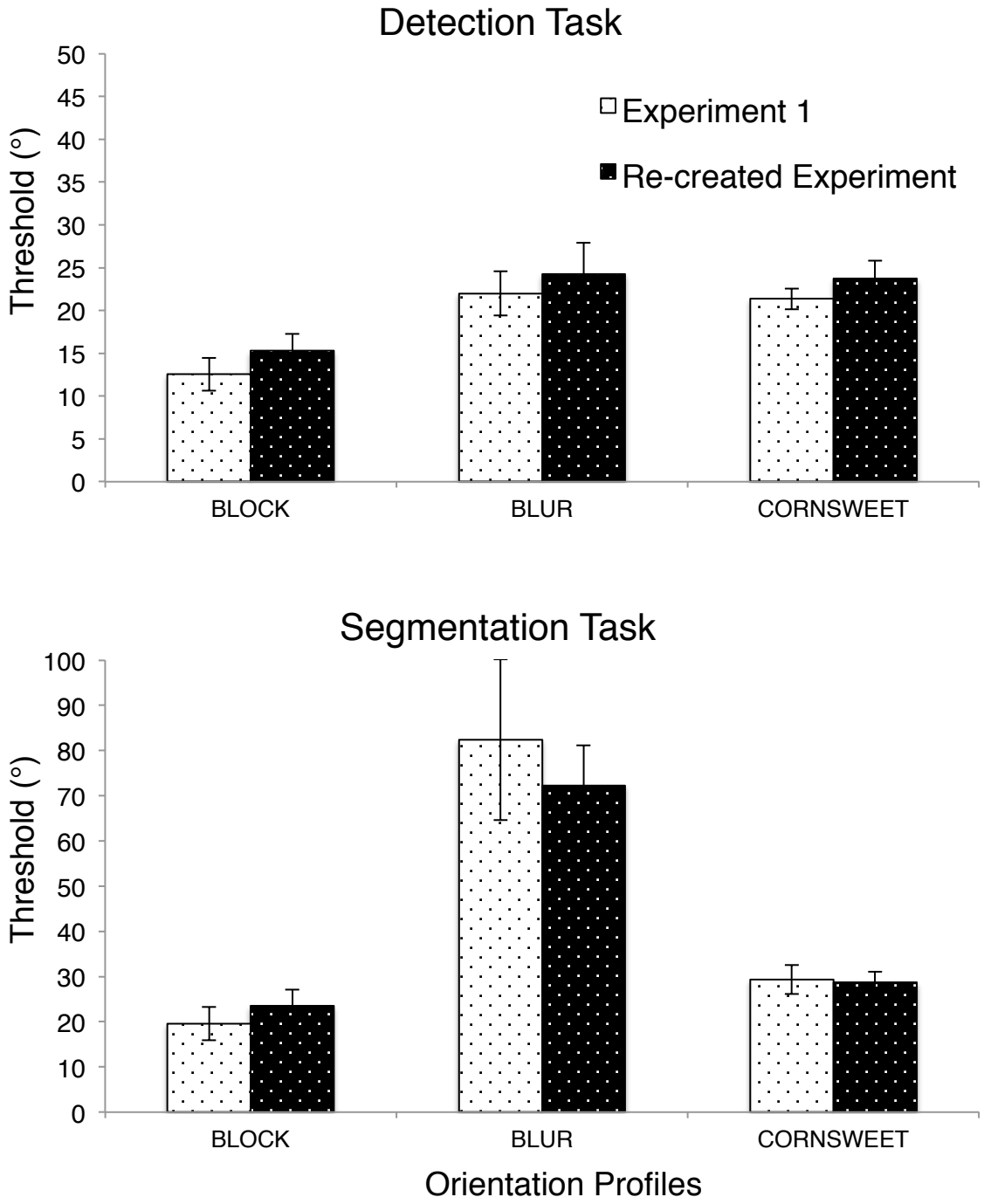


Figure 4.8. Threshold values of participant SKS for Experiment 1 and the Re-created Experiment. Error bars represent the 67% confidence interval. The stimulus duration was 108ms, and orientation jitter of the line elements were 7.5° standard deviation. The Re-created Experiment was produced to be identical to Experiment 1.

4.2 Experiment 2: Effect of Parameter (Steepness) of Orientation Profiles

From the modeling of Experiment 1, the prominent finding is that integration region does not increase with time. A possible interpretation that could account for the result is that texture segregation is mediated by a fixed-size second-order filtering mechanism i.e. the second linear filter of the FRF model. The FRF model is characterised by two linear filters, the first of which uses channels that are orientation-selective and spatial frequency tuned. This first stage filter is useful in detecting luminance-defined patterns, but is unhelpful in detecting texture-defined patterns. The second linear filter on the other hand is orientation tuned, and operates at a **coarser** spatial scale, which results in selectivity for the texture pattern i.e. able to detect texture-defined boundaries (see Section 2.7.1 for more details on the FRF model). The results of Experiment 1 suggest the presence of such a large-scale i.e. coarse second-order mechanism, whose size remains fairly constant.

To test this theory, the current experiment was designed. As with the previous experiment, the Block, Blur, and Cornsweet profiles were used while participants performed the segmentation and detection task. However, for the Blur and Cornsweet profiles, the parameter that determines the curve of the slope which varies the orientation change was manipulated (see Figure 4.9). By changing the parameter, we get to manipulate the degree of orientation change within a particular region. If some type of a fixed-size second-order filtering mechanism was present, threshold performance for the Blur profile should improve with a steeper slope (lower parameter value). However, for the Cornsweet profile, performance should decline with a steeper slope (higher parameter value).

For the current experiment, we also manipulated the positional arrangements of the line elements. The line elements could be positioned uniformly (i.e. in a grid), or randomly (both with overlap and without overlap). The purpose of this manipulation was to investigate if

collinearity was in some confounding manner influencing the outcome of the experiment. As we discussed in Section 2.3 of the introduction, collinearity amongst texture elements can improve texture segregation (e.g. Giora & Casco, 2007; Harrison & Keeble, 2008). When line elements are arranged in grid (as in Experiment 1), the alignment of the elements by the figure border could produce collinear effects that have nothing to do with texture segregation. Thus, by using random placement of line elements, we can ensure that the outcome we obtain is purely due to texture segregation, not the other configural effects.

4.2.1 Methods

4.2.1.1. Participants

Four observers (DK, SKS, SA, and NA) participated in the experiment. DK and SKS were experienced psychophysical observers, while SA and NA were paid volunteers who were naïve observers. All participants had corrected-to-normal vision.

4.2.1.2. Stimuli

The same three orientation profiles – Block, Blur, Cornsweet – from Experiment 1 were used (see Section 3.3 for more details). However, in this study, the parameters that control the steepness of the slope (k and f respectively for the Blur and Cornsweet profiles, see Equations 1-3 in section 3.3) were manipulated. The Blur parameter values used were 22.5 [same as Experiment 1], 16.875, 11.25, and 5.625, while the Cornsweet parameters used were 0.026180 [same as Experiment 1], 0.03927, 0.05236, and 0.06545 (see Figure 4.9 for illustration). Decreasing the Blur parameters resulted in a steeper slope (Figure 4.10), while increasing the Cornsweet parameter resulted in a steeper slope (Figure 4.11).

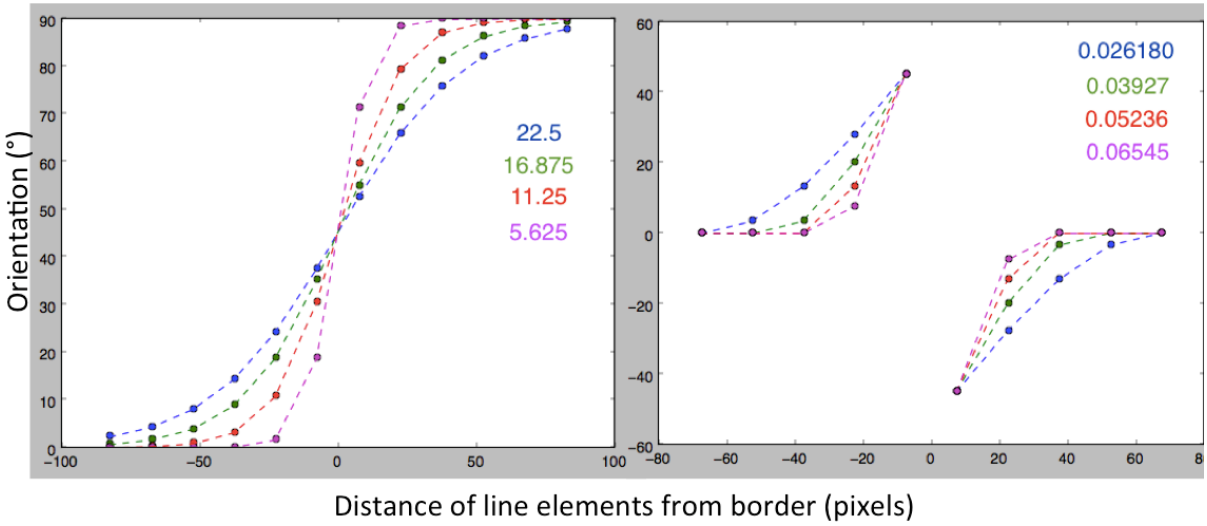


Figure 4.9. The steepness of the slope changes with the different parameters (Blue, Green, Red, and Purple) of the Blur (left) and Cornsweet (right) profile. The y-axis is the orientation of the line element, while the x-axis is the distance of the line elements from the border. The border is represented as 0 here, with positive values referring to line elements positioned within the figure border, and negative values are line elements positioned outside the figure border.

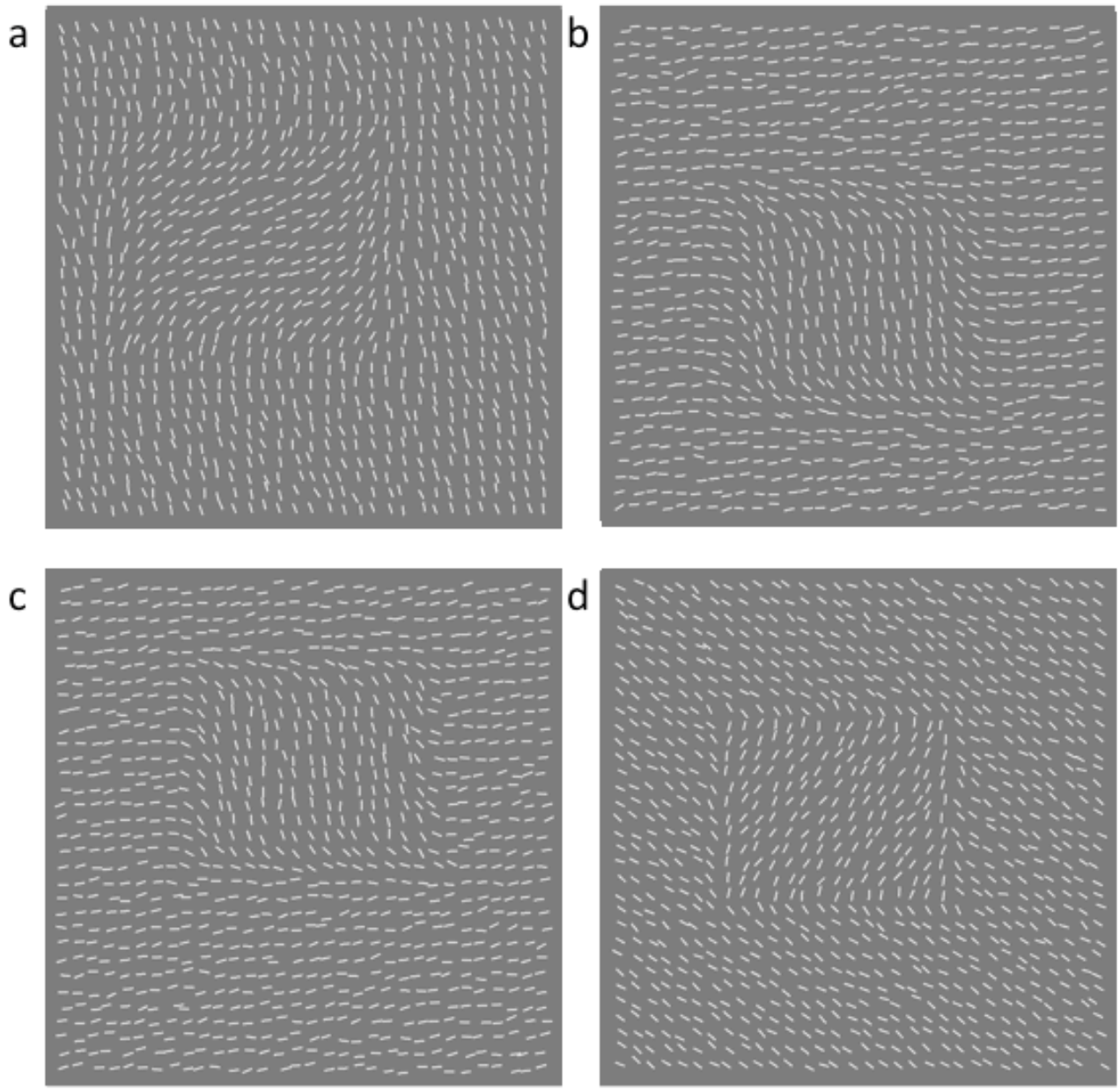


Figure 4.10. Examples of stimuli for the Blur profile with parameter values of (a) 22.5, (b) 16.875, (c) 11.25, and (d) 5.625. Decreasing the parameter value results in a steeper slope of the logistic curve, which in turn makes the orientation change more rapidly near the border of the figure. This results in a figure that segregates more easily. For all 4 stimuli shown here, the orientation contrast is set at 90°, with 7.5° Standard Deviation orientation jitter.

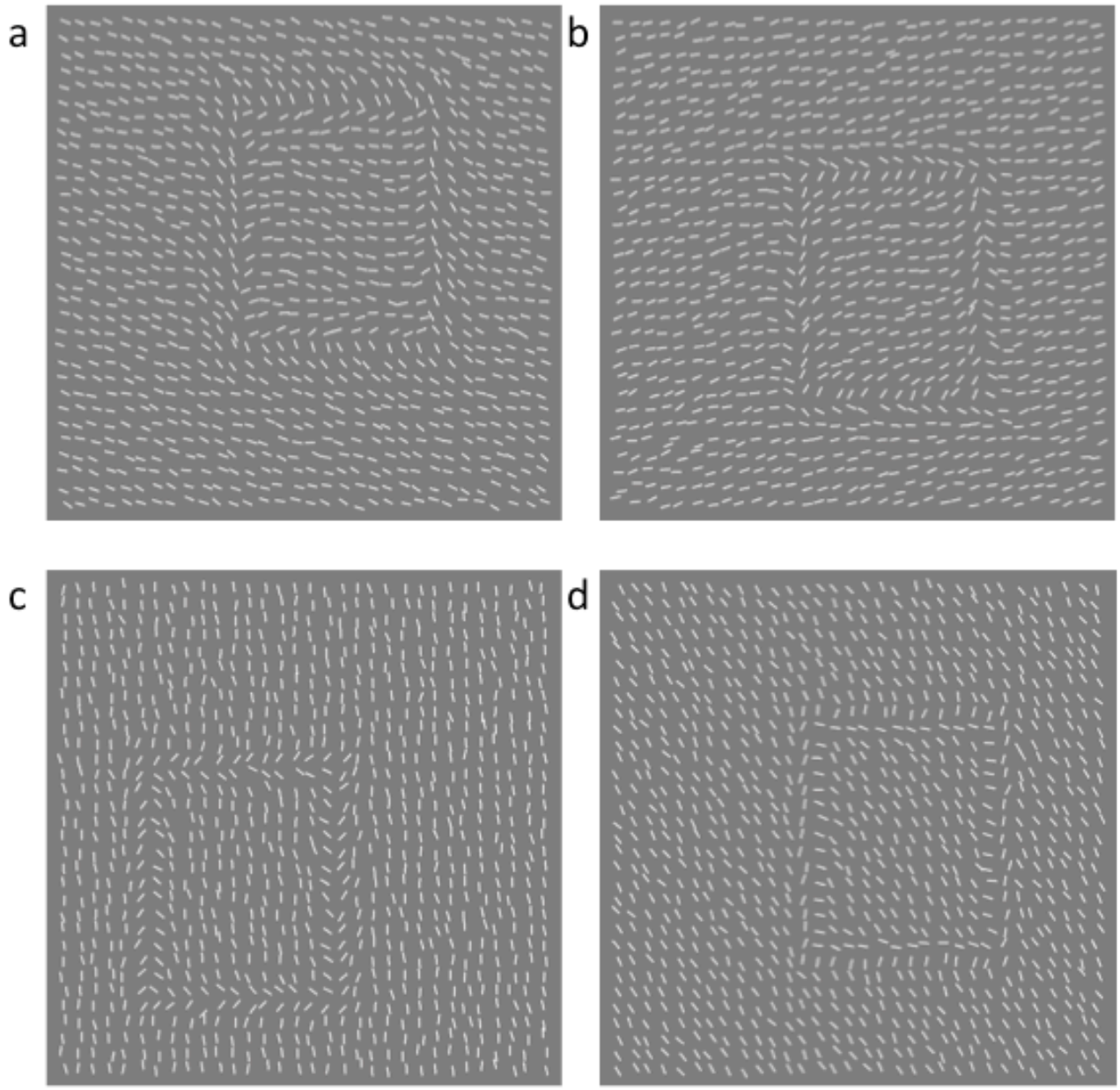


Figure 4.11. Examples of stimuli for the Cornsweet profile with parameter values of (a) 0.026180, (b) 0.03927, (c) 0.05236, and (d) 0.06545. Increasing the parameter value results in a steeper slope of the Cornsweet curve, which in turn makes the orientation change more rapidly near either side of the border of the figure. This results in a figure that does not segregate easily (especially at low orientation contrast). For all 4 stimuli shown here, the orientation contrast is set at 90°, with 7.5° Standard Deviation.

The placements of the line elements were also varied so that they took on three different types of arrangements – Grid [standard condition, as described in Section 3.3], Random, and Exclusion (Figure 4.12). For all arrangements, the line elements appeared on an 8.0°×8.0° area. For the Random arrangement, 1024 line elements were placed randomly within the 8.0°×8.0° area. The number of line elements for both the Grid and Random arrangements were the same. The Exclusion arrangement also had random placements of the line elements, but unlike the Random arrangement, there was no overlap of the line elements. There were 720 line elements in the Exclusion arrangement, with a distance of at least 2 pixels between the ends of the lines.

The stimuli for the Exclusion arrangement were constructed as follows:

1) If the number of elements is less than 720, keep finding a position to place the element
2) Chose a position randomly within the 480×480 pixel area
3) Check whether the ends of the new line element are too close (i.e. less than 2) to previously placed line elements
4) If it is too close, the program returns to 2. If it is not too close, the element is placed

The reason why we had to use fewer elements for the Exclusion arrangement is that as the number of elements increase, the program gets caught in an infinite loop going through steps 2 – 4, where the terminating condition can never be met.

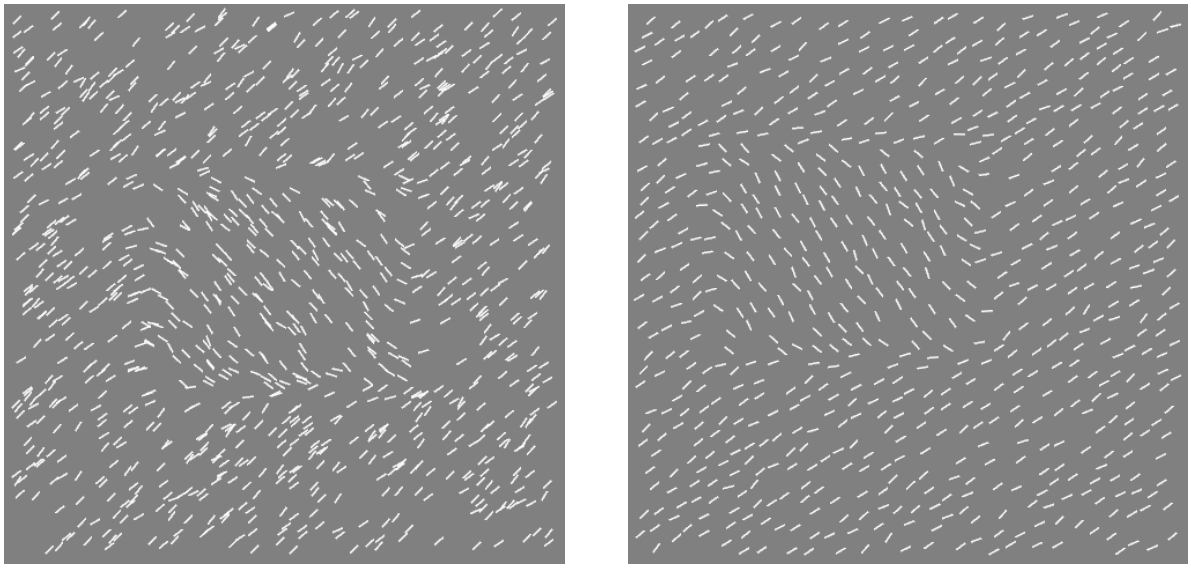


Figure 4.12. Example of stimuli with Random (left), and Exclusion (right) positional arrangements.

For all three arrangements, the size of the figure region (within the border) was 225×180 pixels (3.8°×3.0°) for the horizontal figure and 180×225 pixels (3.0°×3.8°) for the vertical figure. This translated to a 15×12 and 12×15 figure for the Grid arrangement, and an average of 180 and 127 line elements for the Random and Exclusion arrangement. Furthermore, the border of the figure region would always be at least 67.5 pixels (or 5 lines for the Grid arrangement) away from the edge of the stimulus array. The orientation jitter was set to 7.5° standard deviation.

4.2.1.3. Procedure

In this study, participants performed both the Segmentation and Detection task (see Section 3.4). The method of constant stimuli was used. During any given block of 180 trials, only one task (segmentation/detection) and one arrangement (grid/random/exclusion) were presented. In each block, the different profiles and parameters (9 in total) were presented in random order. Each block was repeated 10 times, with a total of 60 blocks run per participant over multiple days.

4.2.2 Results

Threshold values were in terms of orientation contrast at the border of the figure for the Block and Cornsweet profiles, whereas for the Blur profile, thresholds were in terms of orientation contrast between the central region of the figure and the ground region. Thresholds for the different task and positional arrangement are plotted in Figures 4.13 – 4.20, with Figures 4.19 and 4.20 showing the mean values across participants.

4.2.2.1 Effect of Orientation Profile and Parameter

The trends of the results are in line with what we hypothesised would occur if texture segregation is mediated by a second-order filtering mechanism that had a fixed size. A decrease in parameter (steeper slope) for the Blur profile produces lower thresholds. That is, as the parameter of the Blur profile decreases from 22.5 to 5.625, the average difference in thresholds between the Block and Blur profile decreases from 30.3% to 12.6%. In fact, in most cases, the steepest slope not surprisingly has thresholds very similar to the Block profile. For the Cornsweet profile, an increase in parameter (steeper slope) increases the thresholds of the Cornsweet profile by 9.9% on average However, very rarely do the Cornsweet profile thresholds ever match the Block profile.

4.2.2.2 Effect of Task and Positional Arrangement

The trends in the data also show that performance is overall approximately 53% poorer with the Segmentation Task compared to the Detection Task. Finally, performance for the Grid arrangement was best, followed by the Exclusion arrangement (an average increase in threshold by 14.6%), and worse performance for the Random arrangement (an average increase in threshold by 44.5% compared to the Grid arrangement).

Further inspection shows that for the segmentation task alone, performance was better by 58.0% for the grid arrangement compared to the random arrangement. For the exclusion arrangement, the experienced psychophysical observers (DK & SKS) performed similarly to the grid arrangement, with only a difference in threshold of 12.2%. The naïve observers (AS & NA) on the other hand, had thresholds for the exclusion arrangement that were 27.1% lower than the grid arrangement. For the detection task, the experienced observers performed the same for all the arrangements, with differences in thresholds between the arrangements ranging from 1.5% to 6.0%. However, the naïve observers' thresholds were 45.3% higher for the grid compared to the random arrangement, while the exclusion arrangement had thresholds that were 38.2% higher than the random arrangement. The average difference in thresholds between the exclusion and grid arrangement was only 4.8%.

Most importantly, for all positional arrangements, an increase in parameter for the Cornsweet profile results in worse performance, while a decrease in parameter for the blur profile results in better performance. Therefore, the manipulation of the positional arrangement did not influence the outcome of the orientation profiles or parameter. This implies that our findings are based on texture segregation, and not other facilitative effects due to collinearity.

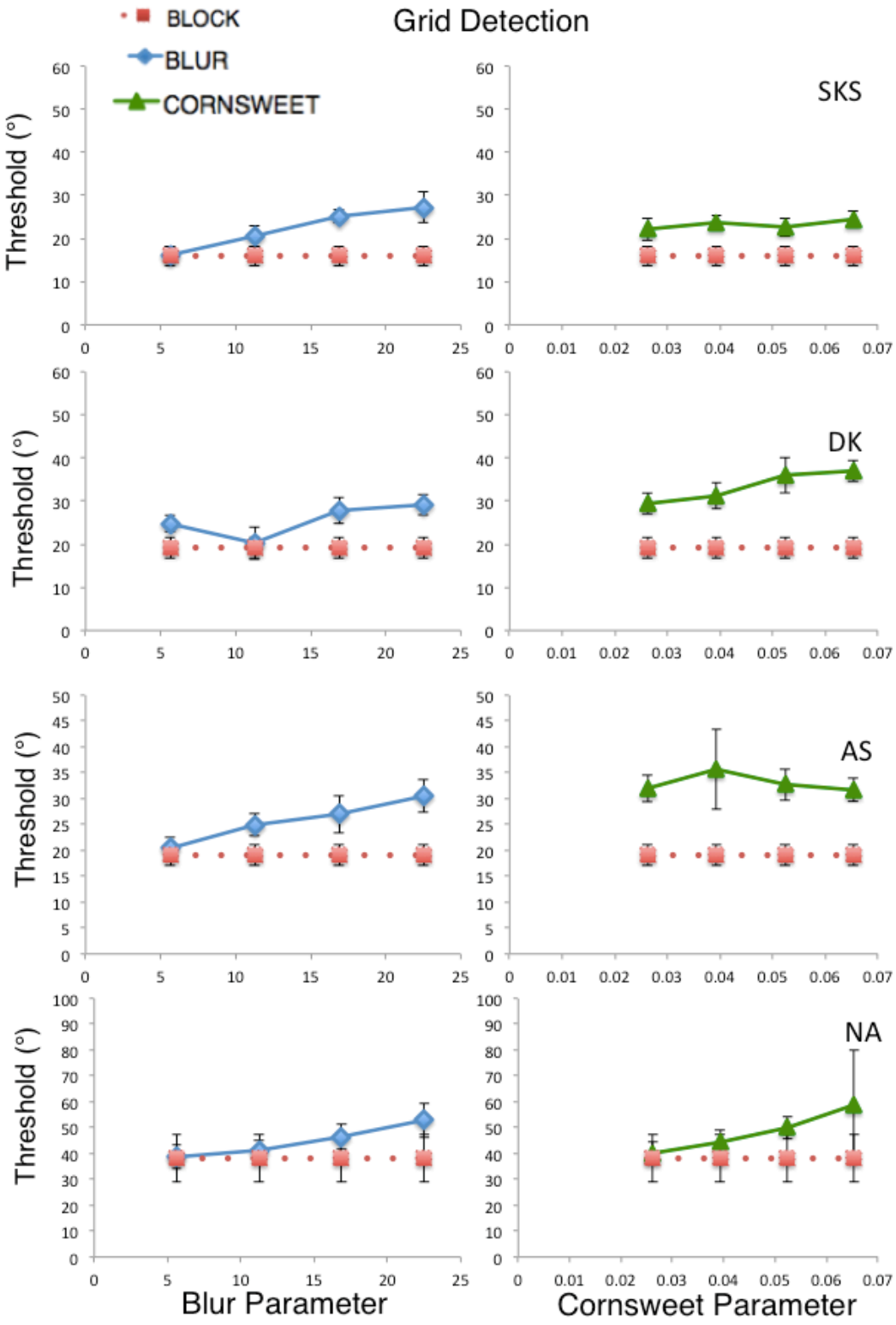


Figure 4.13. Data plots for the Detection Task (grid arrangement) that shows the threshold of the participants. The Blur parameter (left column) and Cornsweet parameter (right column) are represented on the x-axis, and the Block profile is the dashed line (which does not have a parameter change). Error bars represent the 67% confidence interval.

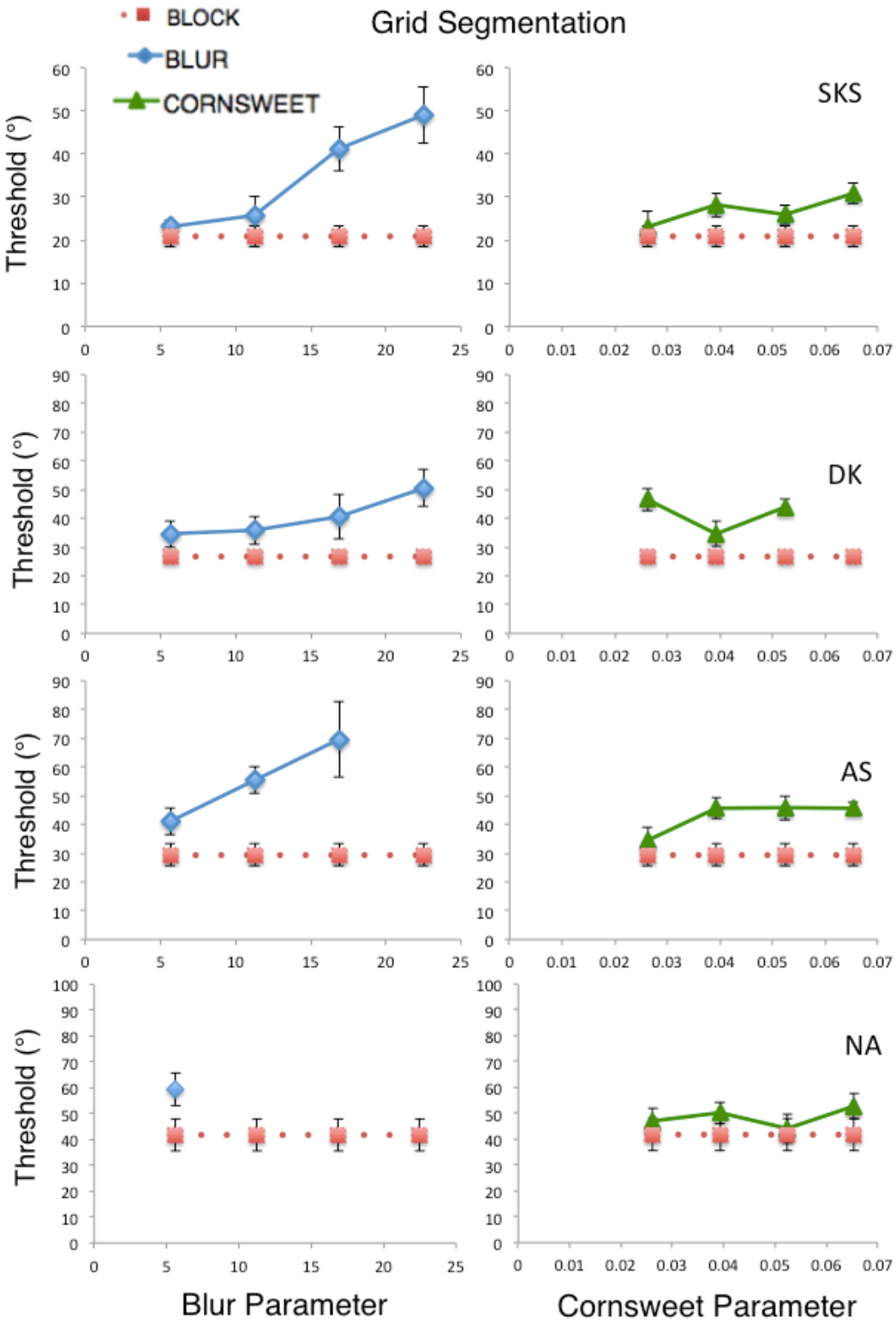


Figure 4.14. Data plots for the Segmentation Task (grid arrangement) that shows the threshold of the participants. The Blur parameter (left column) and Cornsweet parameter (right column) are represented on the x-axis, and the Block profile is the dashed line (which does not have a parameter change). Error bars represent the 67% confidence interval.

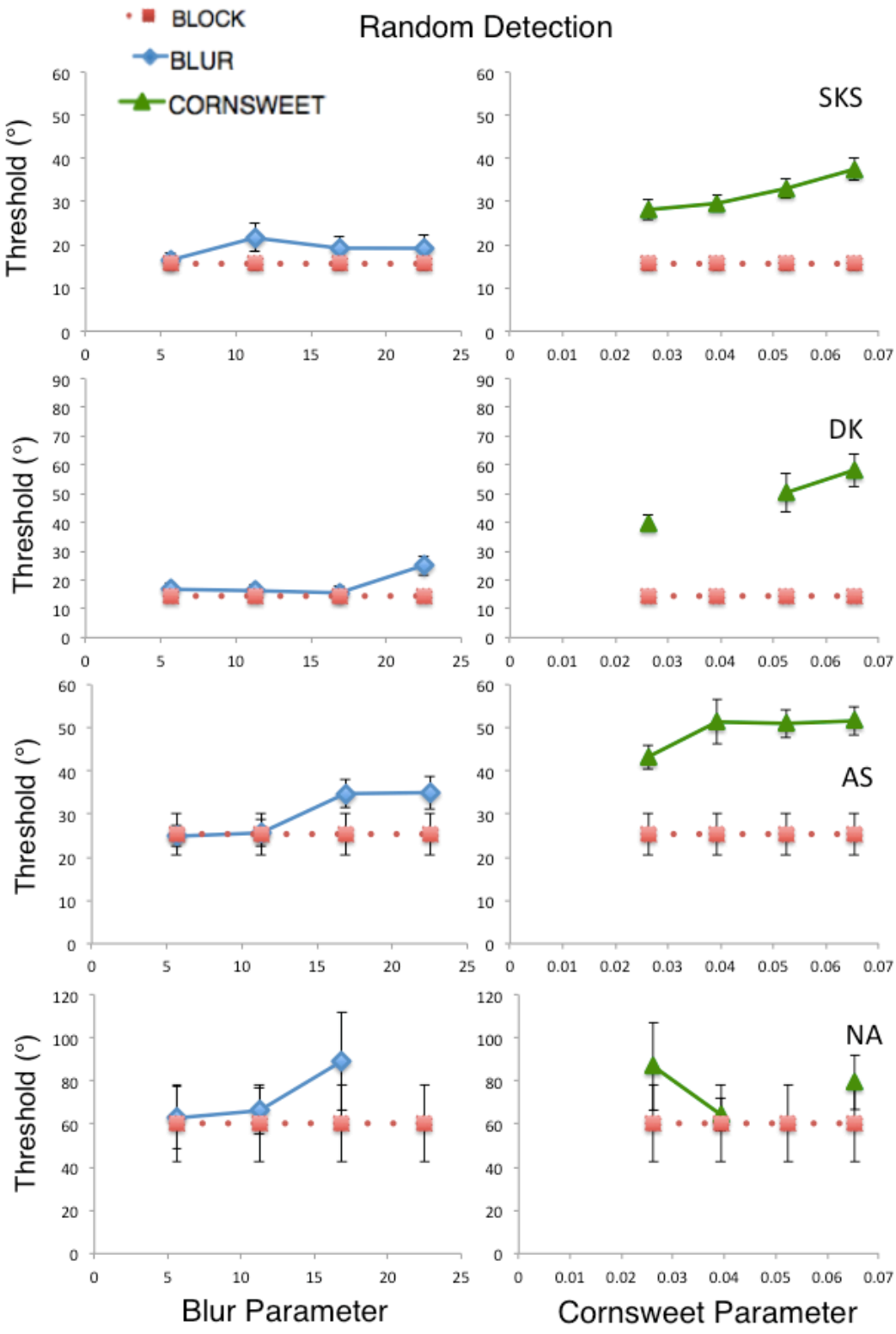


Figure 4.15. Data plots for the Detection Task (random arrangement) that shows the threshold of the participants. The Blur parameter (left column) and Cornsweet parameter (right column) are represented on the x-axis, and the Block profile is the dashed line (which does not have a parameter change). Error bars represent the 67% confidence interval.

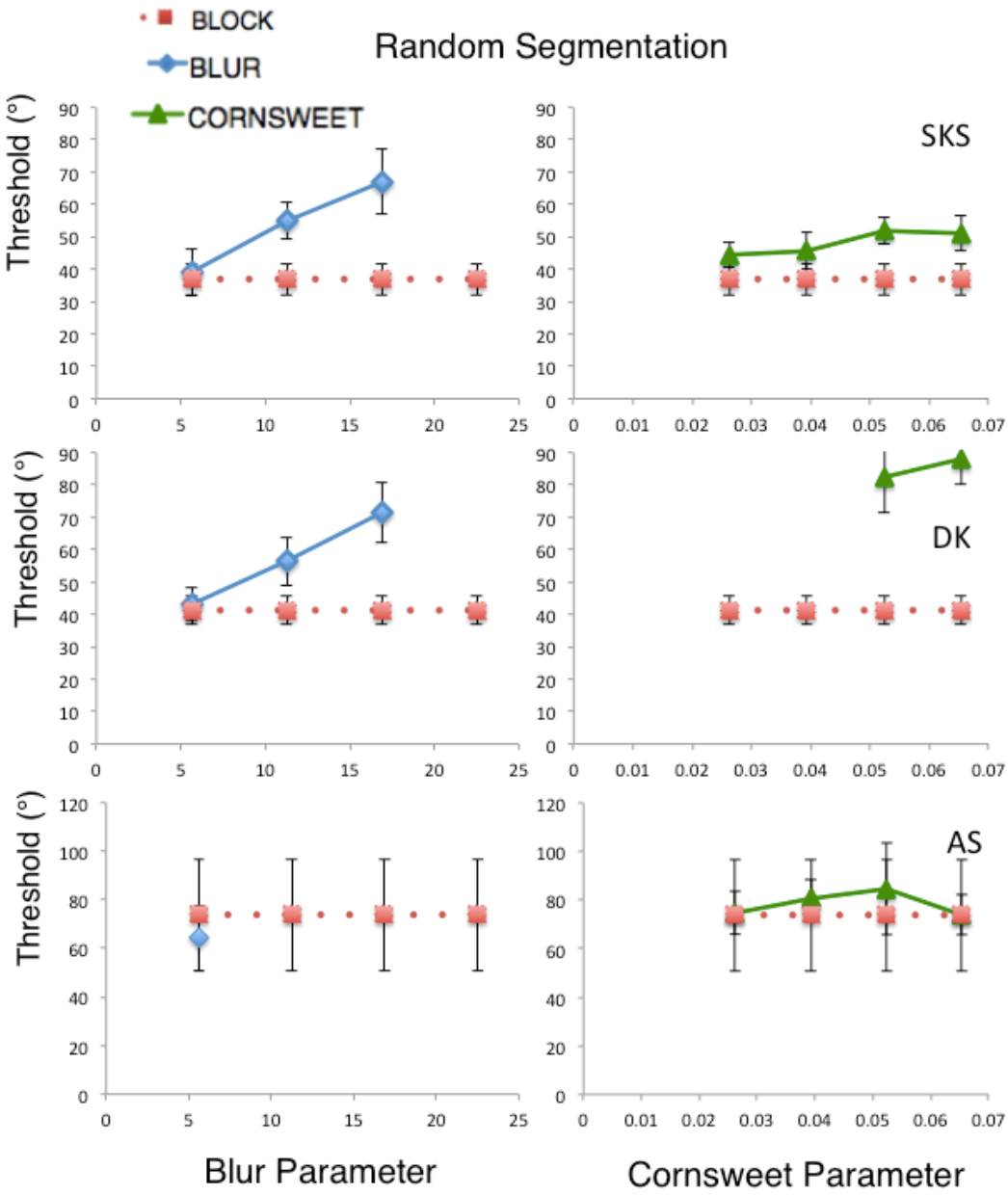


Figure 4.16. Data plots for the Segmentation Task (random arrangement) that shows the threshold of the participants. The Blur parameter (left column) and Cornsweet parameter (right column) are represented on the x-axis, and the Block profile is the dashed line (which does not have a parameter change). Participant NA could not produce physically meaningful thresholds (i.e. <90°) for this task and positional arrangement. Error bars represent the 67% confidence interval.

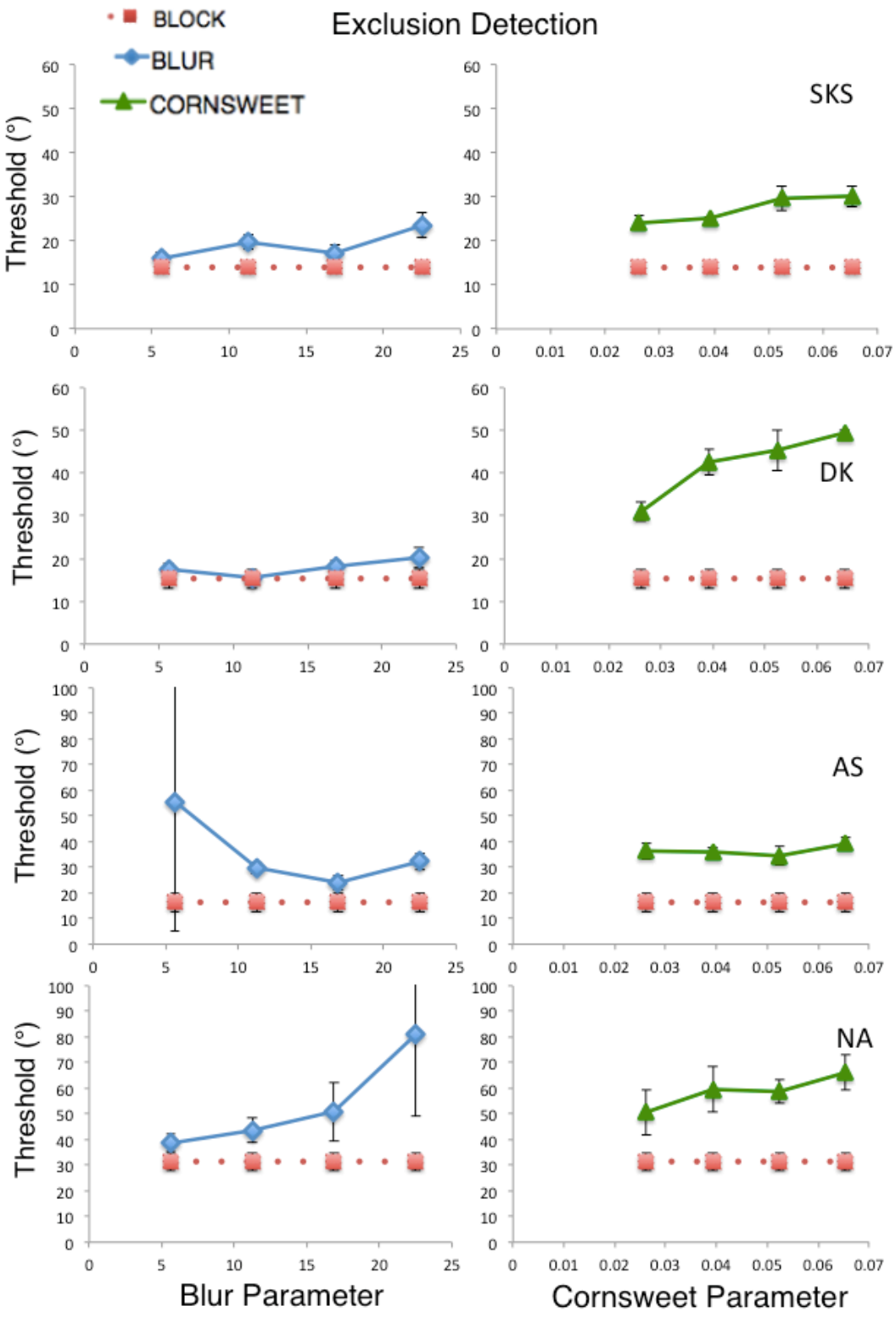


Figure 4.17. Data plots for the Detection Task (exclusion arrangement) that shows the threshold of the participants. The Blur parameter (left column) and Cornsweet parameter (right column) are represented on the x-axis, and the Block profile is the dashed line (which does not have a parameter change). Error bars represent the 67% confidence interval.

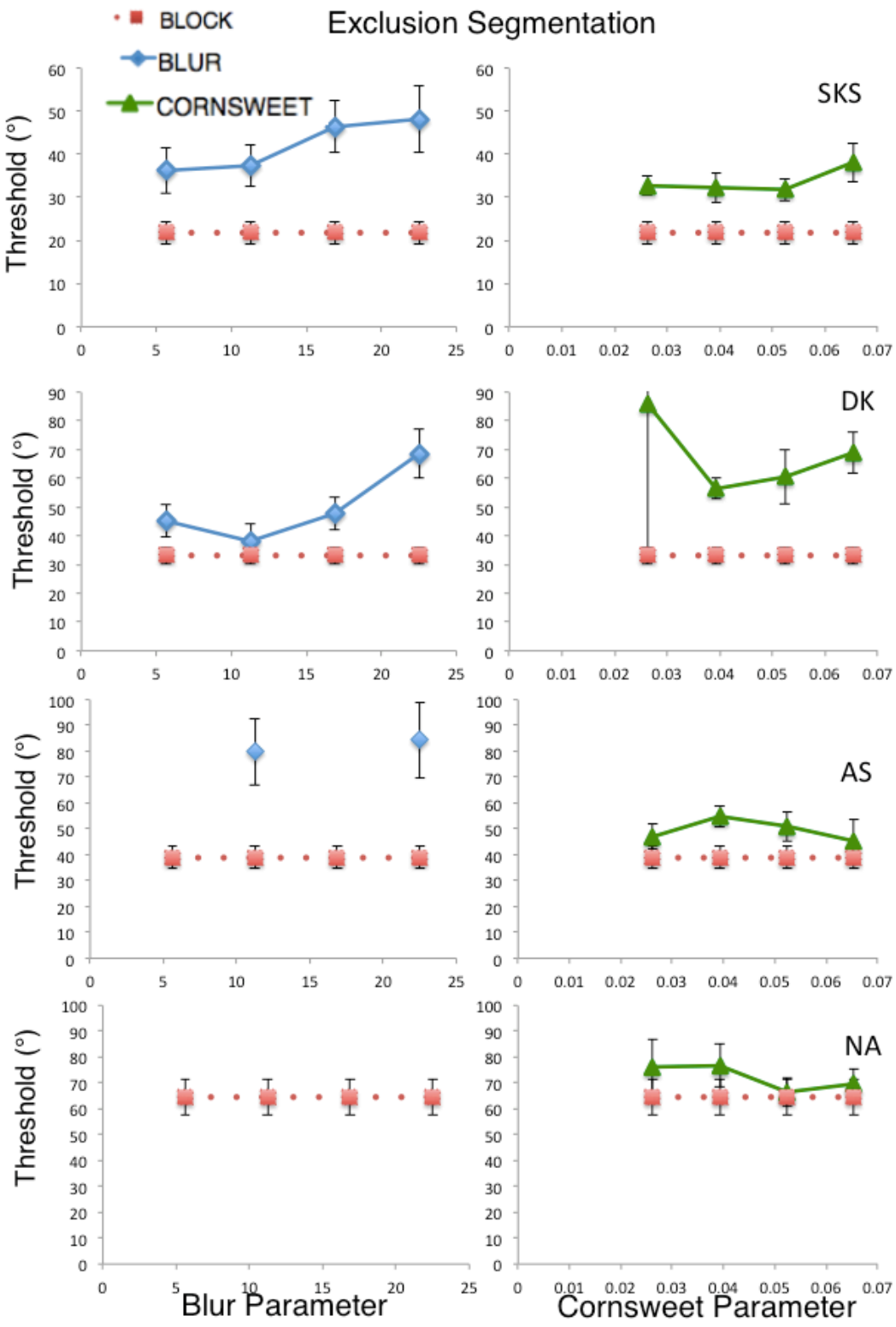


Figure 4.18. Data plots for the Segmentation Task (exclusion arrangement) that shows the threshold of the participants. The Blur parameter (left column) and Cornsweet parameter (right column) are represented on the x-axis, and the Block profile is the dashed line (which does not have a parameter change). Error bars represent the 67% confidence interval.

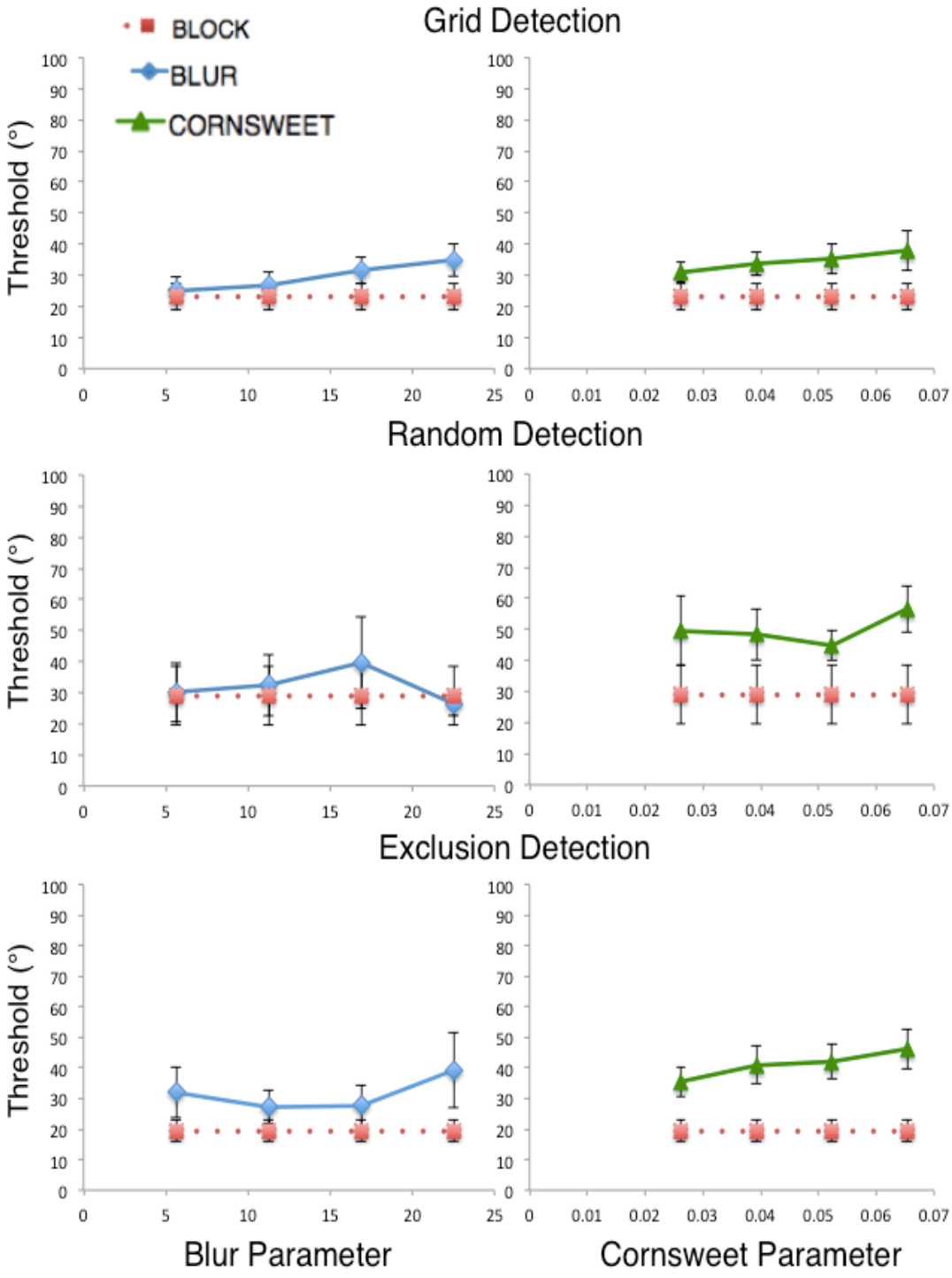


Figure 4.19. Average thresholds for all participants (n=4) for the Detection Task. The different positional arrangements are plotted on the different rows. The Blur parameter (left column) and Cornsweet parameter (right column) are represented on the x-axis, and the Block profile is the dashed line (which does not have a parameter change). Error bars represent the standard error of the mean.

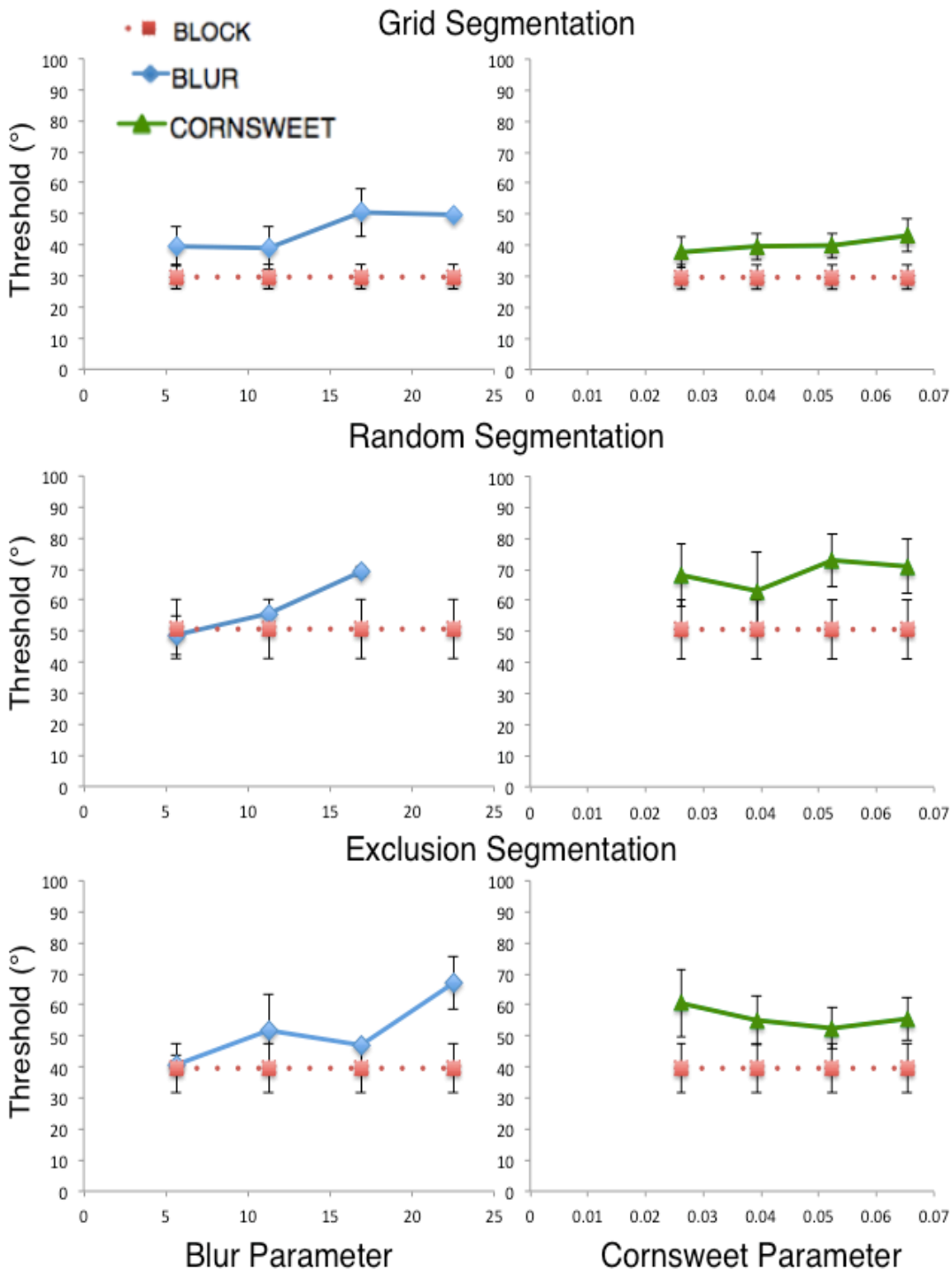


Figure 4.20. Average thresholds for all participants (n=4) for the Segmentation Task. The different positional arrangements are plotted on the different rows. The Blur parameter (left column) and Cornsweet parameter (right column) are represented on the x-axis, and the Block profile is the dashed line (which does not have a parameter change). Error bars represent the standard error of the mean.

4.2.3 Modeling

The modeling of the IR was done in the same manner as described in Section 3.5. However, for the modeling of this study, the IR was calculated from the thresholds of the Block profile, and all parameters of the Blur and Cornsweet profile (i.e. a total of 9 thresholds were used where possible, see Figure 4.21 for example). If a participant's threshold is above 90° (i.e. not a physically meaningful value as an orientation contrast greater than 90° is not possible), the data was removed from the IR analysis. As long as the participant had at least 1 threshold from each profile, we were able to obtain an IR (see Figure 4.22). IR for the Segmentation Task (Random and Exclusion arrangement) for participant NA could not be modelled as a result of this.

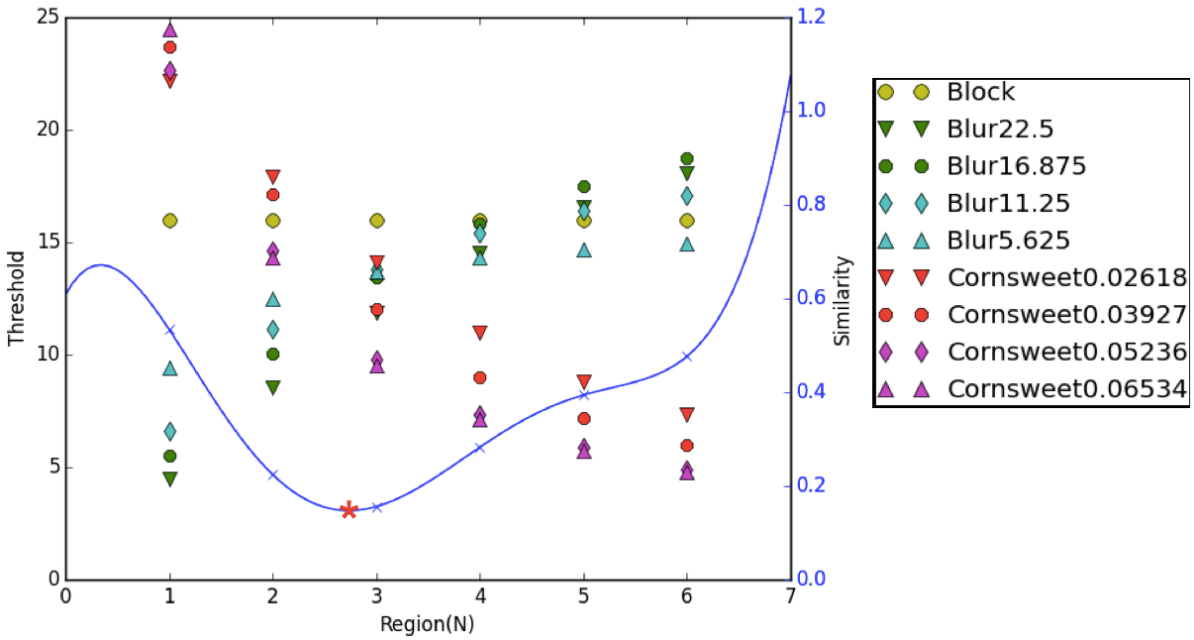


Figure 4.21. An example of how the IR was computed for this study (Participant SKS, Detection Task, Grid arrangement). An integration region was computed from the thresholds of the different profiles and parameters (9 in total). The lowest point on the curve (red *, which represent the point at which the 9 thresholds are most similar) is used as the IR.

The IRs for the four participants (2 tasks: segmentation and detection; 3 positional arrangement: Exclusion, Grid, and Random) are plotted in Figure 4.23. The modeling results show that IR is affected by the task, whereby the IR for the detection task is larger than the IR for the segmentation task. IR also does not appear to be affected much by the positional arrangement (except for participants SKS and DK whose IR for the Detection Task grid arrangement is slightly smaller than the random and exclusion arrangement).

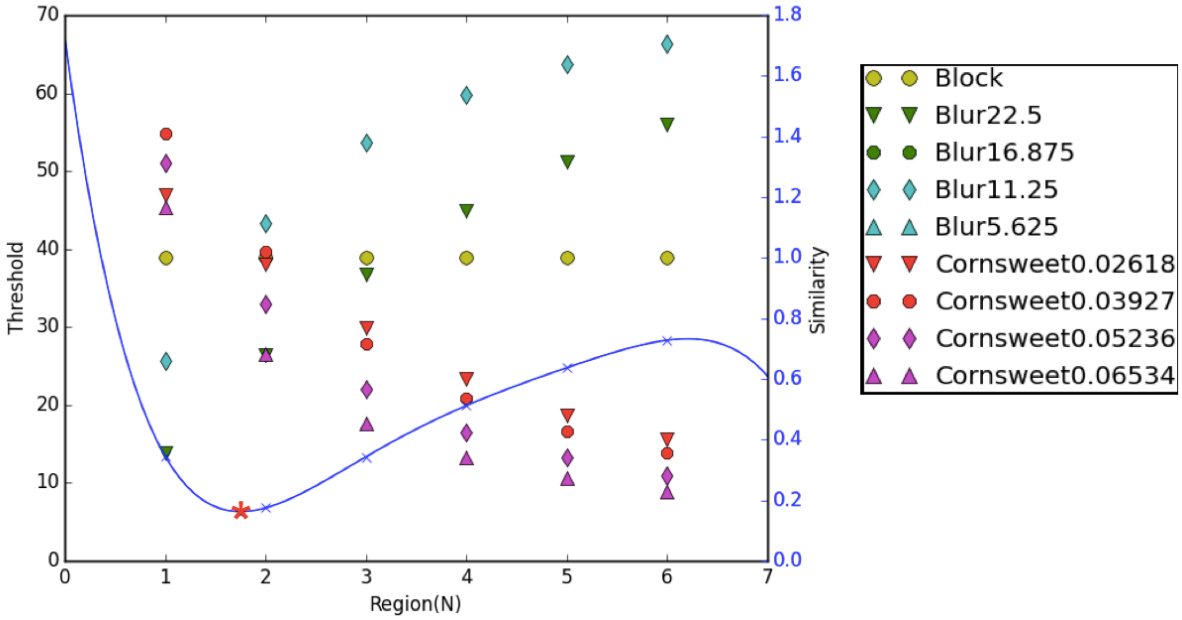
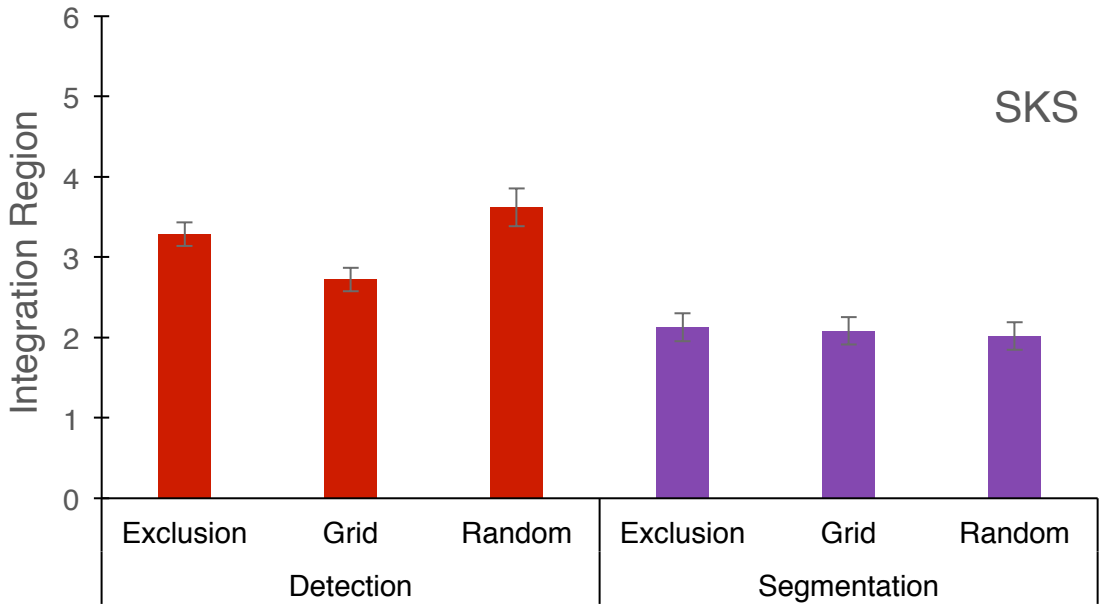
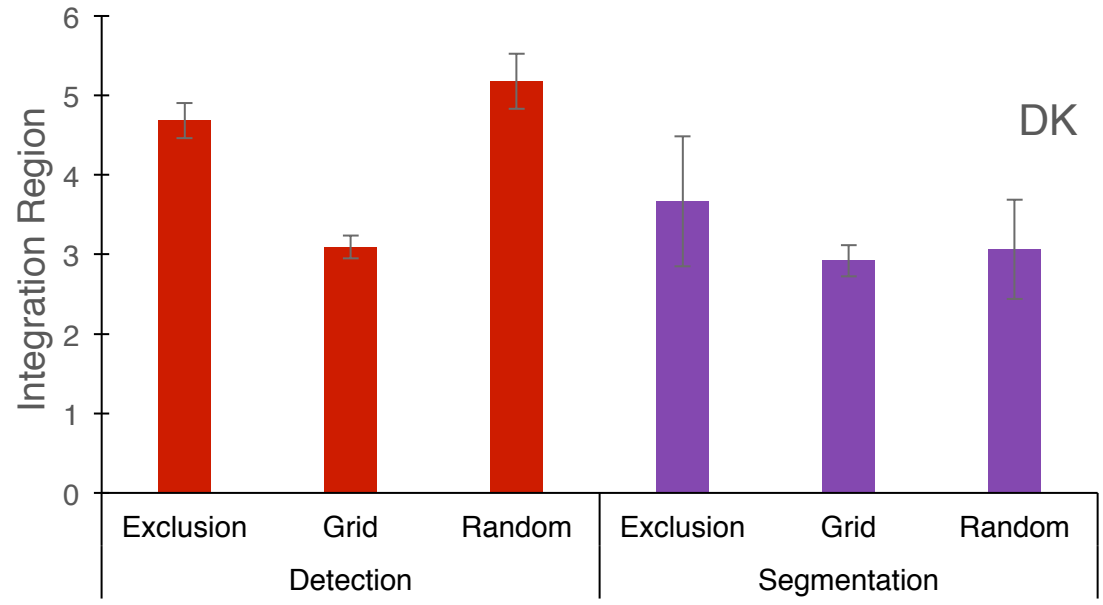


Figure 4.22. An example of how the IR was computed with 7 thresholds (Participant AS, Segmentation Task, Exclusion arrangement). The lowest point on the curve (red *, which represent the point at which the 7 thresholds are most similar) is used as the IR.


SKS
Integration Region
6
5
4
3
2
1
0
Exclusion
Grid
Random
Exclusion
Grid
Random
Detection
Segmentation


DK
Integration Region
6
5
4
3
2
1
0
Exclusion
Grid
Random
Exclusion
Grid
Random
Detection
Segmentation

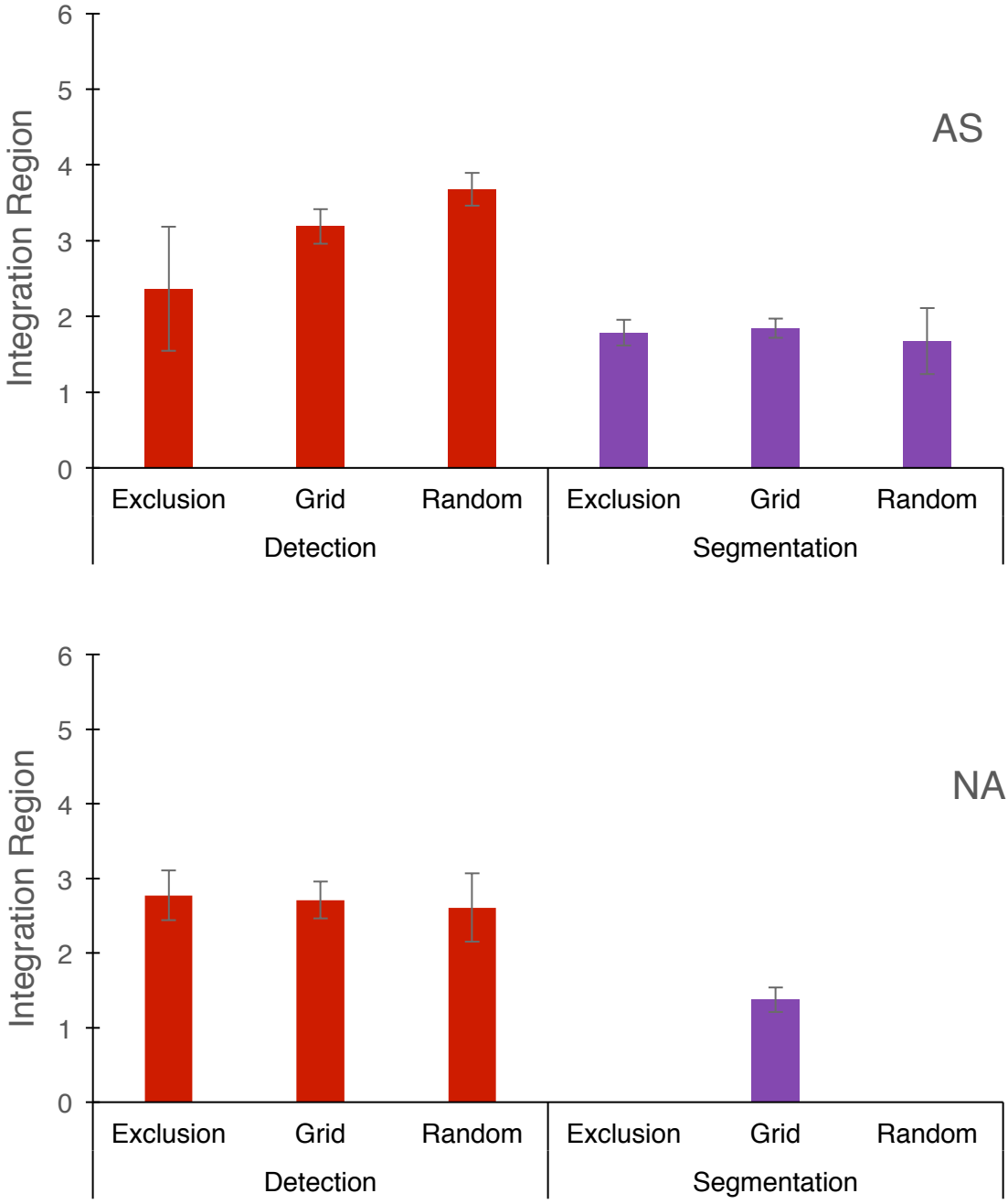


Figure 4.23. Integration Region of the four participants for the two tasks and three positional arrangements. Error bars are the standard deviation, which were obtained by bootstrapping the data.

4.3 Discussion

The purpose of Experiment 1 was to evaluate the effect different types of orientation profiles have on texture segregation. The outcome of the study shows that: 1) hierarchy of threshold

can be expressed as: Blur-centre > Cornsweet > Block > Blur-edge, 2) segmentation task had higher thresholds of orientation contrast compared to the detection task, 3) the increase in orientation thresholds for the segmentation task compared to the detection task is highest for the Blur profile compared to the Block and Cornsweet profile, 4) increased orientation jitter resulted in increased orientation thresholds, and 5) increased stimulus duration resulted in lower orientation thresholds.

In this experiment, texture segregation was studied in two tasks – segmentation and detection. Segmentation refers to the process in which an edge boundary is detected between regions of different orientation contrast. Segmentation is particularly useful for discriminating shapes, as it requires the edge boundary to be explicitly formed. Detection on the other hand is the ability to identify the presence of two distinctive texture regions in the absence of an explicit texture boundary (e.g. Wolfson & Landy, 1998). It has been suggested that segmentation relies on an edge-based mechanism that is able to detect a boundary of orientation contrast, as an observer would need to perceive an explicit edge to perform the segmentation task, while a distinction between two regions would suffice for a detection process (e.g. Wolfson & Landy, 1998; Landy & Graham, 2004; Norman, Heywood & Kentridge, 2011; Vancleef et al., 2013).

The results show that orientation thresholds for the segmentation task are higher than the detection task for all three orientation profiles. However, for the Blur profile, the increase in orientation threshold for the segmentation task is considerably higher compared to the increase in orientation threshold for the Block and Cornsweet profile. Considering that the segmentation task requires an explicit edge to be formed (e.g. Wolfson & Landy, 1998), segregation that relies on a region-based mechanism would yield poorer performance compared to an edge-based mechanism. This could account for the greater decline in

performance for the Blur profile, as a region-based mechanism is most likely operating in the absence of an abutting edge. For the Block and Cornsweet profile, an explicit border can be formed via an edge-based mechanism, and thus performance declines only moderately for the segmentation task compared to the detection task.

We initially posited that if texture segregation occurs due to an edge-based mechanism, thresholds for the Block and Cornsweet profile would be similar, as both profiles have information of orientation contrast at the edge of the figure. In contrast, if texture segregation were due to a region-based mechanism, we would expect that the thresholds for the Block and Blur profile would be similar to each other, as both have information of orientation contrast at the centre of the figure.

Our results show that in most cases, the Block and Cornsweet profiles did not have similar threshold values, which suggests that segregation is not occurring due to information at the edge-region only. In fact, the extremely low thresholds of the Blur profile as a function of edge orientation contrast (Blur-edge) is not very likely, as the presence of an edge-based mechanism that detects a discontinuity at the exact figure border should yield thresholds that were similar for the Block, Cornsweet, and Blur-edge. As such, for further experiments, we will not be analysing the Blur thresholds as a function of orientation contrast at the edge.

Furthermore, performance for Block profile was also better compared to Blur-centre even though the information within the central most region of the figure was the same for both profiles. This result suggests that texture segregation is not entirely due to a region-based mechanism either. Our claim is that the Block profile, which can be detected by both an edge and region-based mechanism, would segregate better compared to the other profiles that rely on a single mechanism to segregate the regions.

Our results show that performance for the Blur profile, as a function of orientation contrast between figure centre and ground, was poorer compared to the Block and Cornsweet profile, especially in the segmentation task. This finding is corroborated by studies that show poorer performances when orientation change occurs over a larger distance (e.g. Nothdurft, 1985b; Landy & Bergen, 1991). However, Kingdom and Keeble (1996) found that performance in terms of amplitude of orientation modulation for a smoothly-varying sine-wave (comparable to a Blur profile) while worse than a square-wave (comparable to a Block profile), was better than a missing-fundamental wave (comparable to a Cornsweet profile). We estimated from graphs, the average orientation thresholds of all participants in their study: thresholds were 2.76° for the square-wave, 3.57° for the sine-wave, and 10.75° for the missing-fundamental. Comparable averages of detection thresholds for the current study were 17.24° for the Block, 24.83° for the Cornsweet, and 28.62° for the Blur profile. While the task used by us and Kingdom and Keeble was the same (2-interval forced-choice detection task), they used Gabor patches instead of line elements, and the orientation jitter was with a standard deviation (SD) of 10°, while the SD for the compared results used for our study was for 7.5°. Also, the stimuli used in the studies, while having a similar pattern in orientation change, are different in terms of composition (see Figure 2.15 in Chapter 2 for Kingdom and Keeble (1996)'s stimuli, and Figure 3.3 in Chapter 3 for our stimuli). While a direct comparison cannot be made due to the differences between the studies, the orientation thresholds from Kingdom and Keeble (1996) were much lower than the current study. Moreover, while performance for the Block (square-wave) was the best for both experiments, in the current study, performance for the Blur profile was marginally poorer than the Cornsweet profile, which is in contrast to what Kingdom and Keeble (1996) found for their comparable stimuli.

However, our results do show that it is possible to segregate a Blur figure from the background. Since it is possible to perform segregation, it suggests the presence of two

mechanisms for texture segregation, one that extracts information from the edge, and also another that extracts information from regions other than the edge. The latter mechanism is most likely responsible for the segregation of smoothly varying stimuli like the Blur profile. This is in accordance with studies that show that smoothly-varying stimuli are able to segregate (e.g. Ben-Shahar & Zucker, 2004). However, segregation of smoothly-varying stimuli is said to be due to the disruption to curvatures found within the stimulus (Ben-Shahar & Zucker, 2004; Ben-Shahar, 2006; Ben-Yosef & Ben-Shahar, 2008). A key point to note about this is that these curvatures are a result of a frame field representation, which is a frame field of the normal and tangential vectors (see Figure 2.16d in Chapter 2). Therefore these take into account the position and orientation of each line elements as it varies smoothly, and the physical representation in degrees of orientation contrast by itself is not able to predict curvatures i.e. stimuli with 90° orientation contrast could have different curvatures based on the baseline orientation of the elements, and the change in orientation of the line elements from background to figure. Thus, in our study where the baseline orientation i.e. background orientation of the line elements was randomly selected for every trial, we have not been able to compute the curvatures present in the Blur profile. Therefore, it cannot be said either way if these curvatures reliably produce segregation.

Another prediction we had was in regards to the filling-in mechanism (e.g. Caputo, 1998; Lamme et al., 1999). If edge extraction precedes filling-in, we expect that performance for the Block profile would be better compared to the Cornsweet profile at longer stimulus durations owing to the **additional** orientation contrast information present at the centre of the figure. However, at short stimulus durations, performance for the Block and Cornsweet profiles would be similar at short stimulus durations when filling-in would not yet have occurred

According to Lamme et al., (1999) who found evidence of a filling-in mechanism, processing of an orientation based texture stimuli begins with local feature detection, which leads to detection of feature boundaries, and finally a representation of figure and background surfaces. Neuronal responses to the stimulus begin at 20-30ms after stimulus onset, and figure-ground responses only begin at 70-80ms (with strongest responses at receptive fields near the boundary). At 90-100ms, boundary responses continue to increase while no changes have occurred for figure centre. At 115-125ms, peak responses to boundaries are observed, with only a slight response to figure centre. Responses for the whole figure surface (boundary + figure centre) only appear at 150-160ms. These findings supported the notion that boundary/edge detection precedes the process of figure surface filling. Similarly, Poort et al., (2012) found enhanced neuronal responses to the edge of a figure ~35ms before responses to the central-region of the figure. However, they found that surface filling-in is dependent on attention, and would not occur if the stimuli were not attended to.

A key point to note about the aforementioned studies is that they were conducted using single-cell recordings of monkeys. Scholte et al., (2008), on the other hand, used human participants, and likewise found that boundary detection preceded surface segregation i.e. the filling-in of the surface region. They found that activation, which reflected boundary detection, began 92ms after stimulus onset. This started at the occipital area, and moved in a feedforward manner to the peri-occipital area (104ms), temporal area (104-108ms), and parietal area (104-120ms). Contrarily, the first instance of surface segregation activation started in the temporal area 112ms after stimulus onset. Subsequently, activation spread to the peri-occipital area (140ms), parietal area (172ms), and occipital area (172ms).

Based on these studies, it is expected that performance for the Block profile would be better than the Cornsweet profile at longer stimulus durations due to the additional orientation

contrast information at the centre i.e. after surface filling-in has occurred. Taking into consideration that both our stimuli and model are reflective of cells in the visual areas (occipital area), neuronal activity that signals surface segregation in the occipital area should occur at 172ms (Scholte et al., 2008). That is, participants should have improved performance for the Block profile compared to the Cornsweet profile at stimulus durations of 216ms. Our results show that irrespective of time, the Block profile generally had enhanced performance compared to the Cornsweet profile. Thus, our experimental results do not show evidence of a filling-in mechanism. Additionally, the results of our modeling show that the size of the integration region remains relatively constant with increased stimulus duration. Thus, the results do not support the notion of a spreading/filling-in process for texture segregation. According to Scholte et al., (2008), the IR should increase at the 216ms stimulus duration, owing to the surface filling-in. However, we do not find this pattern of results.

Unlike Lamme et al.,(1999), and Poort et al., (2012), who found neuronal enhancement to the figure simply because the receptive field was on the figure region, Rossi, Desimone, and Ungerleider (2001) found that figure modulation (enhanced neuronal response to the figure) was due to orientation contrast specifically at the boundary (i.e. not necessarily due to the figure, but proximity to the boundary). According to Rossi and colleagues, enhanced response to the figure was due to proximity between figure and background, in which contextual influences by elements along the boundary signalled the enhanced response. According to them, enhanced responses to the figure will occur when classical-receptive-field (CRF) is near the boundary region (~1.5°). For the Block profile, an enhanced response is achieved by the figure due to the contextual effects (background and figure elements) by the boundary. For the Cornsweet profile on the other hand, the contextual influences include neighbouring elements away from the edge (for our stimuli, 1.5° encompasses at least 6 rows of lines), which has orientation contrast that is decreasing, which could result in the slightly

poorer performance for the Cornsweet profile. For the Blur profile, orientation contrast is maximal at the centre of the figure compared to the background, which is not spatially close enough to elicit an enhanced response.

According to Li's V1 model (e.g. Li, 1999a) of texture segregation, elements on either side of the border would receive enhanced responses as they experience weaker iso-orientation suppression due to the neighbouring elements that are not iso-oriented (i.e. of its neighbouring elements, only half are of the same orientation). At low orientation contrast (15°), the saliency of elements by the border is not that much stronger compared to the other elements away from the border, and the model does not signal the edge boundary (Li, 2002b). Based on this V1 model, both the Block and Cornsweet profile should experience weak iso-orientation suppression for elements on either side of the border, as these elements have fewer iso-oriented neighbours. Arguably, the Cornsweet profile has elements away from the border with differing orientation (non iso-oriented neighbours). However, away from the edge, the change in orientation for the elements are much smaller compared to at the edge, and according to the V1 model, will not have a strong signal. Despite this, performance for the Cornsweet profile was slightly worse compared to the Block profile. According to the V1 model, the signal strength that would indicate the boundary for both profiles should be the same. For the Blur profile, the change in orientation is constantly low, and based on the model, this would not signal a border. However, our results show that segregation was achieved with the Blur profile. This suggests that segregation is not solely based on signal strength of the border region.

Therefore, even the models proposed by Rossi et al., (2001) or Li (1999a), cannot account for the ability of our visual system to segregate stimuli with smooth texture variations. However, our results show that segregation of smoothly varying stimuli is possible, as in most cases

participants were able to perform the task (i.e. produce physically meaningful thresholds). Thus, to sum up our standing, we have found that neither an edge-based mechanism (that detects a discontinuity at the immediate figure border), nor a region-based mechanism (that groups together at once regions of a texture that are similar) could fully account for our results. Additionally, we have found no evidence of a filling-in mechanism either. We therefore propose that a second-order texture filter mediates the segregation process. The third stage of the FRF model (or the second linear filter) is characterised by its larger size filter, which could explain our results. That is, we propose a mechanism that extracts information over a larger region than the immediate border of the figure. As it is extracting information over a large region, it is akin to a region-based mechanism. However, our claim is that this 'region' is of a fixed size i.e. it does not immediately group together **all** elements that are similar.

In Experiment 2 of this chapter, we tested this theory by varying the parameter that alters the change in orientation of the line elements. Assuming the presence of a fixed-size filter, the different parameters would have different degrees of orientation contrast within a fixed region. We predicted that thresholds would be lower for the Blur profile when the parameter value was lower (steeper slope), and that thresholds for the Cornsweet profile will be higher when the parameter value was higher (steeper slope). The trend in the data was exactly as predicted, suggesting the presence of fixed-size second-order filter.

Moreover, we varied the orientation jitter of the line elements in Experiment 1, we found that the conclusions drawn about the results do not change with the different orientation jitters. We can therefore conclude that the results we have obtained are not limited to the particular orientation jitters we used, but are prevalent findings. However, we did find that performance

generally declines with orientation jitter, and similar effects have been observed by Vancleef et al., (2013) and Li (1999a).

Additionally, we manipulated the positional placements of the line elements in Experiment 2. With the grid arrangement used in Experiment 1, configural effects such as collinearity may well have mediated texture segregation (e.g. Giora & Casco, 2007; Harrison & Keeble, 2008). The key outcome of this manipulation was that the positional arrangements of the line elements did influence our predictions based on the fixed size second-order filter. Moreover, the various positional arrangements of the line elements did not affect the IR. Thus, we can claim that our results are based on the visual system segregating the region purely based on the texture properties within the stimuli.

Vancleef et al., (2013) studied the influence of element arrangement with discrimination and segregation tasks. They found that the arrangement of elements, either random or regular (comparable to the grid arrangement used in the current study), did not influence performance in the discrimination task (comparable to the detection task). However, with the segregation task (comparable to the segmentation task), in which participants had to indicate if the edge was straight or curved, the regularly positioned elements had improved performance compared to the random arrangement. This indicates that spatial arrangement plays a role in shape discrimination, as needed by a segregation task, but not in a discrimination task in which discriminating two regions from each other will suffice.

In the current study, for the segmentation task, performance was better for the grid arrangement than the random arrangement. This is in line with the findings of Vancleef and colleagues. For the exclusion arrangement, the experienced psychophysical observers (DK & SKS) performed similarly to the grid arrangement, but the naïve observers (AS & NA) performed worse than the grid arrangement, but better than the random arrangement. For the

detection task, for the experienced observers, performance was the same for all the arrangements, but for the naïve observers, performance was better for both the grid and exclusion arrangement. This suggests that the representation of the edge boundary (required to perform a segmentation task) is impaired with the random arrangements, but not for the grid and exclusion arrangement (only for experienced observers). As suggested by Vancleef et al., (2013), this could be because there is a greater variability of distance within the elements with the random arrangement, which leads to an overall drop in the contrast gradient defining the edge. For the exclusion arrangement, the line elements were randomly placed, but did not overlap with each other, and had fewer elements (n=720) compared to the random and grid arrangements (n=1024). According to Nothdurft (1985a), when the spacing between line elements increases, performance will decrease. As a consequence of reducing the number of elements in the exclusion arrangement, the spacing between the line elements would have increased, which could have influenced the naïve observers more than the experienced observers.

Finally, for both the modeling done in Experiments 1 and 2, the IR of the detection task is greater than the IR of the segmentation task by approximately 1 raster unit. The average IR for the detection task was 3.6 [Experiment 1] and 3.3 [Experiment 2], while for the segmentation task it was 2.6 [Experiment 1] and 2.3 [Experiment 2]. A possible explanation of this could be the task itself. As mentioned previously, segmentation requires the edge boundary to be explicitly formed, while detection merely requires the ability to identify the presence of two distinctive texture regions in the absence of an explicit texture boundary (Wolfson & Landy, 1998). Therefore, a larger IR will be useful for the detection task as it will sample a larger portion of the stimulus to detect if there are two distinct regions to the texture. For the segmentation task, a smaller IR will be most efficient in explicitly forming an edge boundary to identify the shape of the region.

To sum up, our results have shown no evidence of an edge-based mechanism, or a filling-in mechanism. We propose a mechanism that is region-based, but not in the sense where it automatically groups together regions of a texture that are similar. The mechanism we propose is that of a large scale second-order filter that has fixed size. In the experiments presented in the following chapters, we investigate what conditions (if any) would affect the size of this filter.

Chapter 5

Evidence for a Fixed Size Second-Order Texture Filter

In Chapter 4, we used stimuli consisting of line elements whose orientations were varied according to different edge orientation profiles to produce figure-ground textures. We found that lower orientation contrast thresholds were produced when the texture figure had information of orientation contrast at both the centre and the edge of the figure, instead of just the centre or edge of the figure. With the assumption that a common mechanism operates to detect and segment textures with the different orientation profiles, we also modelled the results in terms of an integration region (IR). A second-order texture filter of a fixed size best modelled the results of Experiments 1 & 2.

However, studies have shown evidence of multiple second-order texture filters of different sizes (e.g. Sutter, Sperling, & Chubb, 1995; Kingdom & Keeble, 1999). These different filter sizes of the second-stage are tuned to the luminance spatial frequency of the first-stage inputs. That is, high luminance spatial frequency (the first-stage input) would result in a small second-stage filter, while low luminance spatial frequency would result in a bigger second-stage filter (Kingdom & Keeble, 1999). In fact, the modeling results of Experiments 1 & 2 show that the integration region, which is a computational model of the size of the second-order texture filter, is bigger for detecting a texture figure compared to segmenting a texture figure. Hence, for the experiments presented in this chapter, we investigate whether there is a range of second-order filters tuned to different sizes (Experiment 3) and shape (Experiments 4) of a texture figure. For Experiment 3, we measured detection and segmentation performance while we varied the size of texture figure, whereas for Experiment 4 we varied the aspect ratio of the texture figure.

5.1 Experiment 3: Effect of Figure Size

In this experiment, participants performed both the detection and segmentation tasks using stimuli with the three different orientation profiles – Block, Blur, & Cornsweet. However, in this experiment, instead of only using the standard figure size of 12×15 elements, we also used figure sizes of 16×20 and 20×25 elements. Note how with these figure sizes, the aspect ratio (i.e. proportional relationship between width and height) of the figures is kept constant at 4:5. This experiment will allow us to investigate whether the size of second-order texture filter will also become larger when the size of the figure increases.

Additionally, this experiment investigated whether the visual system is able to adjust accordingly to the different inputs of stimuli. That is, we ran participants in blocks where the figure sizes were separated (one block will contain only one figure size) or mixed (one block will contain all figure sizes). If we find that the detection and segmentation of texture figures involve multiple second-order filters tuned to the different sizes for the separated condition, but only a single fixed size second-order filter for the mixed condition, it would imply that the visual system was able to adapt to the size of the figure being presented.

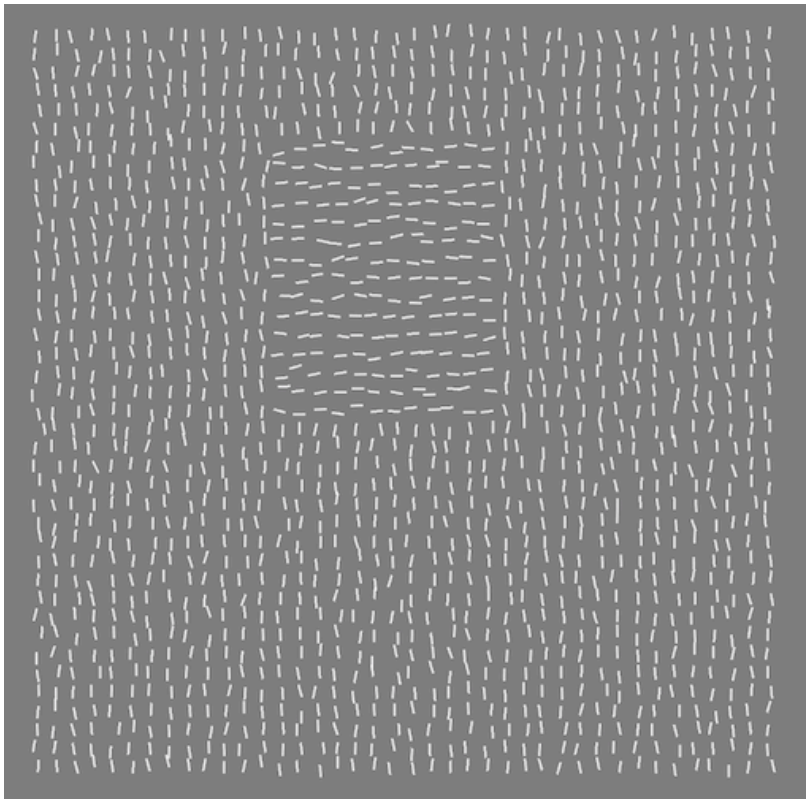
5.1.1 Methods

5.1.1.1 Participants

Three observers (DK, SKS and SA) participated in the experiment. DK and SKS were experienced psychophysical observers, while SA was a naive observer. All participants had corrected-to-normal vision.

5.1.1.2 Stimuli

The stimuli were created as described in Chapter 3, Section 3.3. However, in this experiment, instead of the 32×32 grid, we used a 40×40 grid (10.0°×10.0°). We also used two additional figure sizes – 16×20 (4.0°×5.0°), and 20×25 (5.0°×6.3°) – in addition to the standard 12×15 (3.0°×3.8°) figure size used in the previous experiments (see Figure 5.1). For all three sizes, the aspect ratio of the figure was therefore maintained at 4:5 (or 1.25). All other parameters were the same as described in Section 3.3.



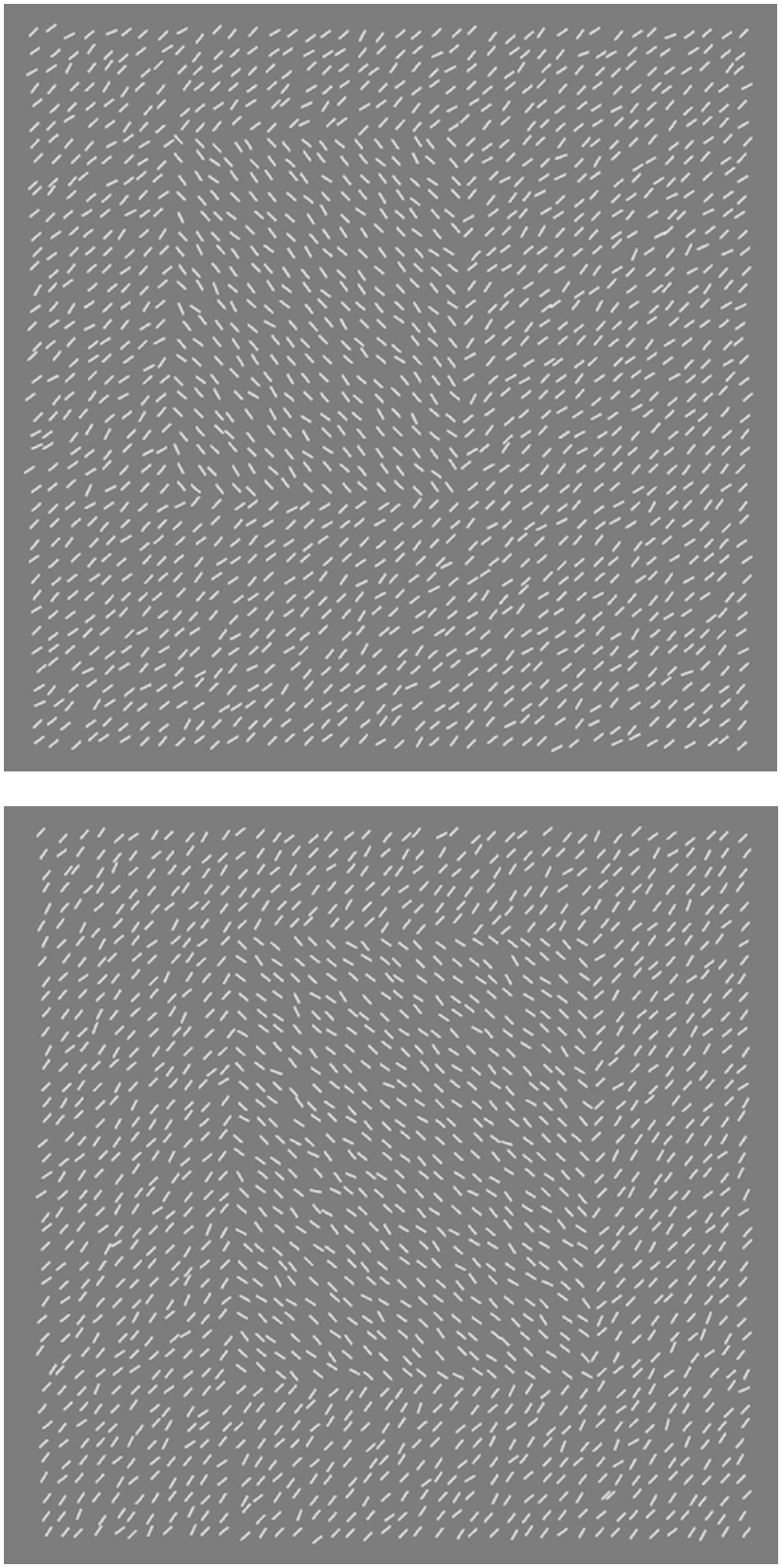


Figure 5.1. Stimuli with the three figure sizes – 12×15 (top), 16×20 (middle), and 20×25 (bottom). All three stimuli are of the Block profile with 90° orientation contrast and 7.5° Standard Deviation for orientation jitter.

5.1.1.3 Procedure

Participants performed both the Segmentation and Detection task (see Section 3.4) as part of this experiment. All three orientation profiles (Block, Blur and Cornsweet) and figure sizes were used. In addition, participants performed the experiment twice – 1) with all the figure sizes mixed within each block (mixed condition) and, 2) with only one figure size within each block (separated condition).

The method of constant stimuli was used (see Section 3.1.1). For the mixed condition, in a given block of 180 trials, all orientation profiles and figure sizes were presented in random order. The different tasks were presented in different blocks. For the separated condition, there were 150 trials per block, with the different orientation profiles presented in random order. The different tasks and figure sizes were presented in different blocks. Each participant ran a total of 44 blocks over multiple days.

5.1.2 Results

Figures 5.2 and 5.3 show the orientation contrast thresholds of individual participants for the separated and mixed conditions respectively, with Figure 5.4 showing the average thresholds across participants. The thresholds do not appear to be influenced by whether or not the figure sizes were separated or mixed together, with the average difference in threshold being only 1.8%. This implies that the visual system does not over time adapt to the size of the figure being presented. Additionally, the thresholds do not vary very much with the different figure sizes tested. However, the exception to this is for the Blur profile of the Segmentation task, where DK and SKS appear to have slightly lower thresholds with larger figure sizes. That is, the difference in thresholds for the different figure sizes were between 1.0% and 2.7% for the Block profile, and between 1.6% and 8.9% for the Cornsweet profile. On the

other hand, average thresholds for the 12×15 Blur profile were 21.6% and 16.6% lower than the 16×20 and 20×25 figure sizes respectively. A similar comparison cannot be made for participant SA as she was unable to produce reliable thresholds for that condition.

We also find the same pattern of results seen in Experiments 1 & 2, where 1) orientation contrast thresholds were on average higher by 26.8% the segmentation task compared to the detection task, and 2) the Block profile had thresholds that were on average 48.0% lower than the Blur profile, and 51.1% lower than the Cornsweet profile.

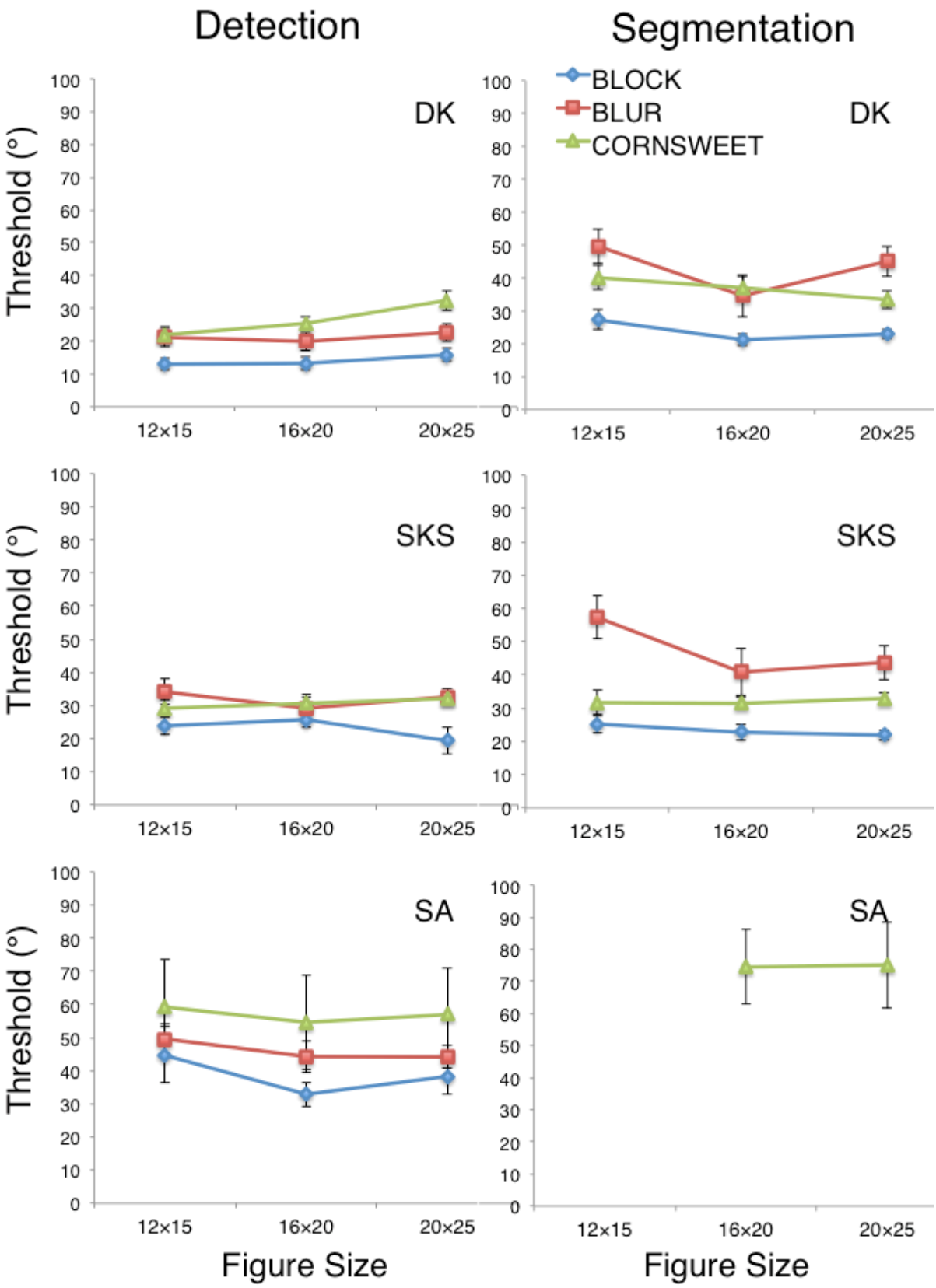


Figure 5.2. Thresholds as a function of figure size for the Detection (left) and Segmentation (right) tasks for the separated condition. Error bars represent the 67% confidence interval.

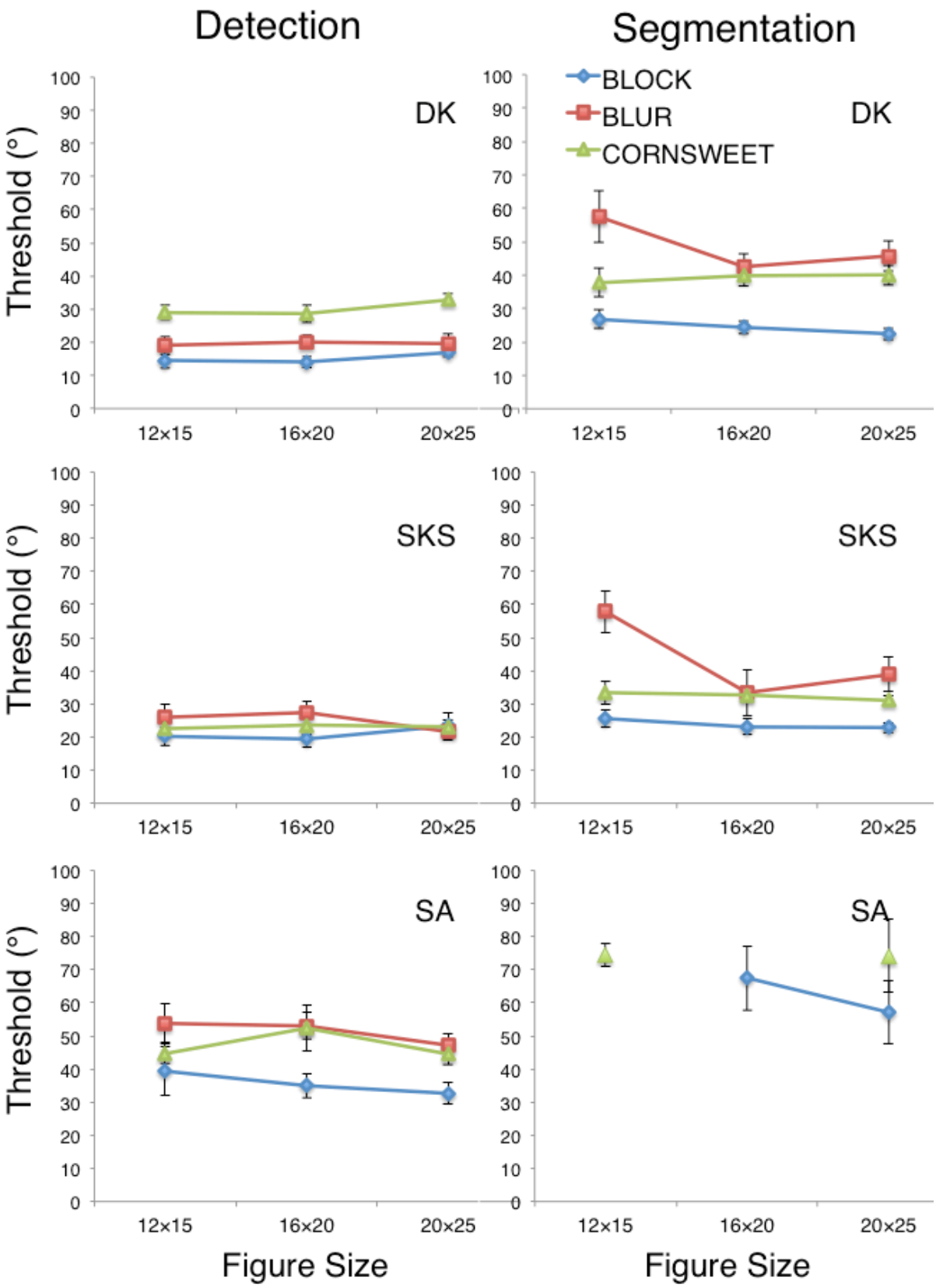


Figure 5.3. Thresholds as a function of figure size for the Detection (left) and Segmentation (right) tasks for the mixed condition. Error bars represent the 67% confidence interval.

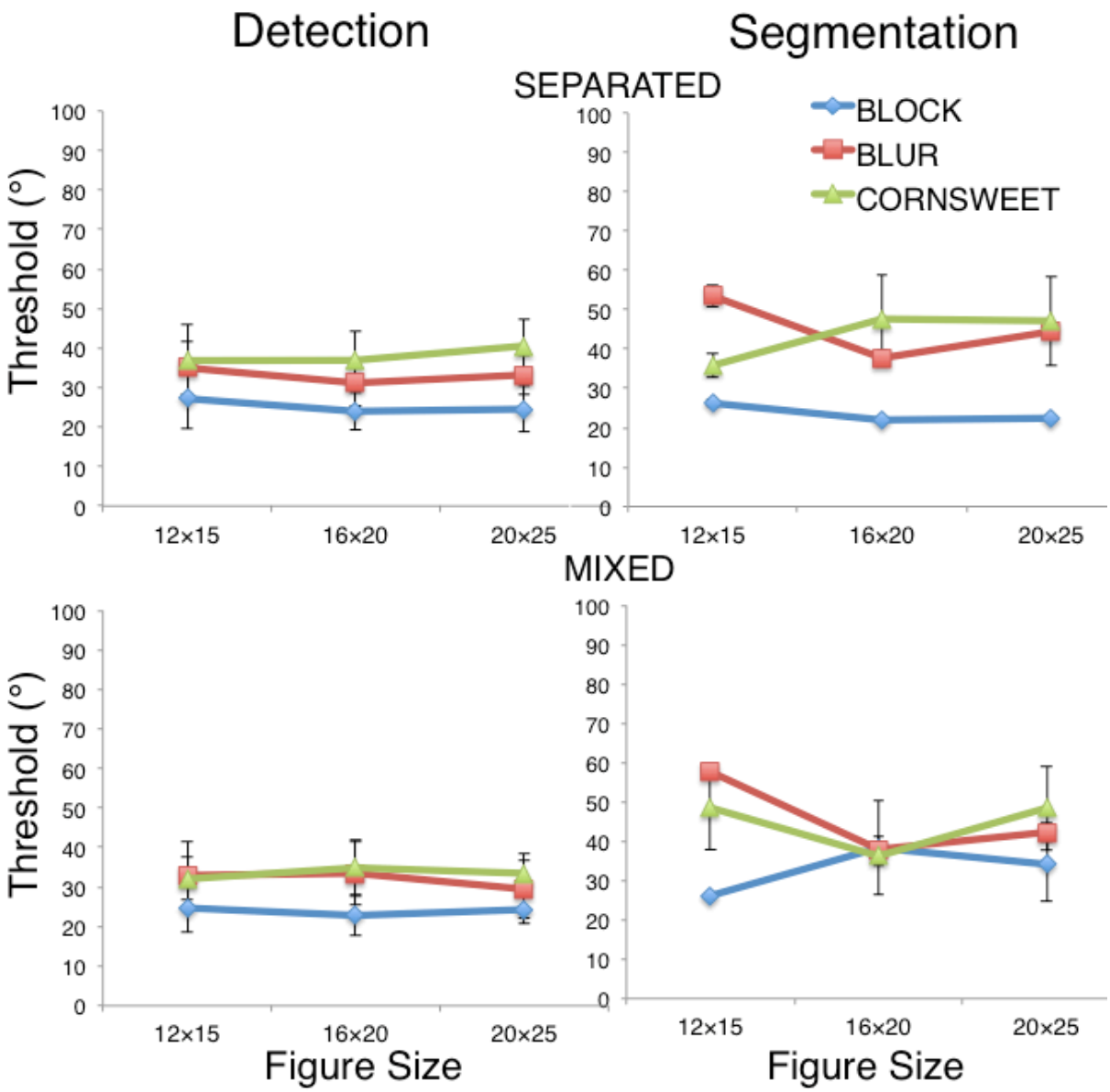
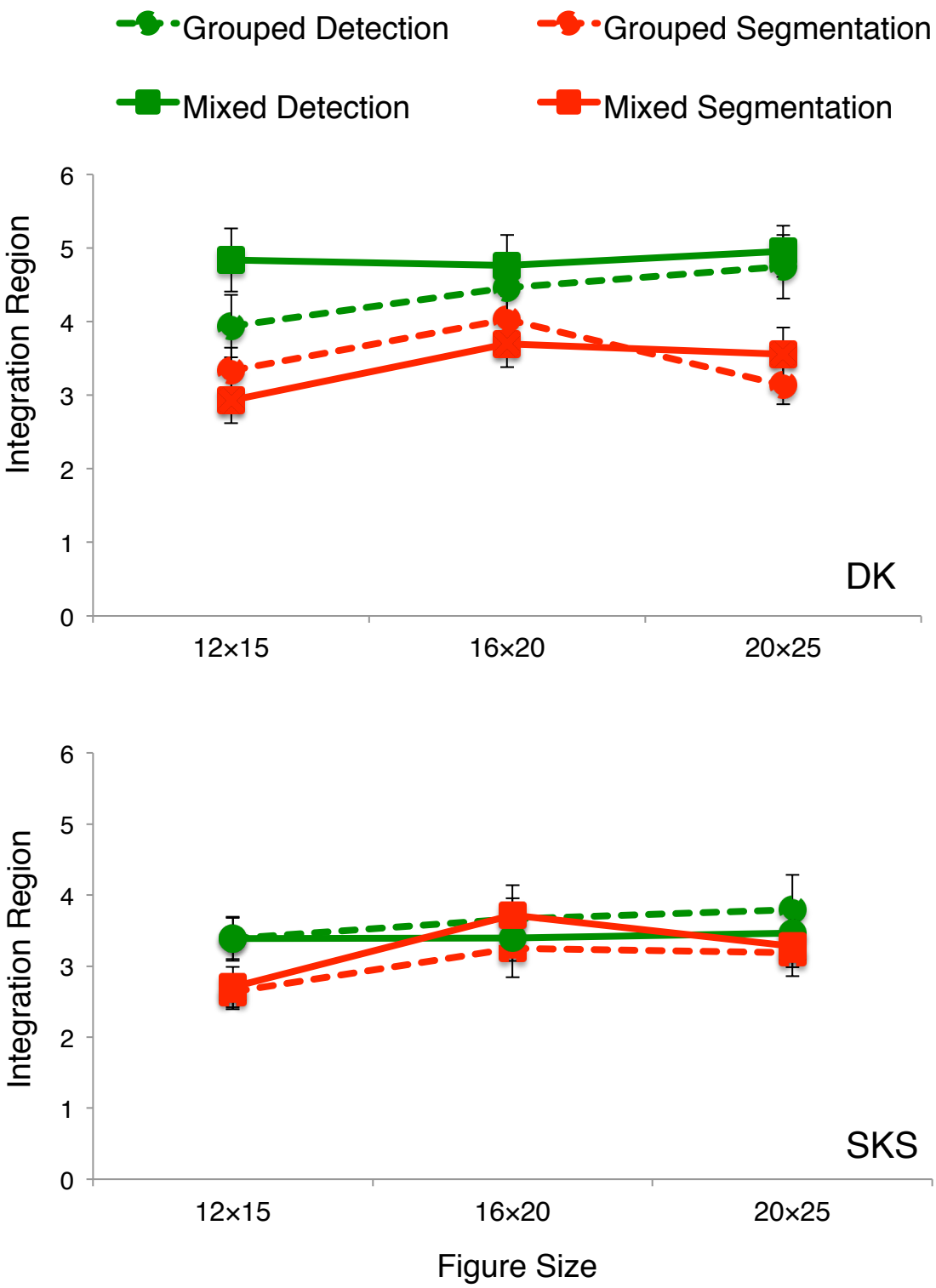


Figure 5.4. Average thresholds of all participants (*n*=3) as a function of figure size for the Detection (left) and Segmentation (right) tasks. The different conditions are plotted on different rows. Error bars represent the standard error of the mean.

5.1.3 Modeling

We model the IR (see Section 3.5) for both tasks and conditions as a function of the figure size. Figure 5.5 shows the IRs for all conditions (2 tasks: segmentation and detection; 2 conditions: separated and mixed; 3 figure sizes: 12×15, 16×20, and 20×25).

The results once again show that the figure sizes do not much influence the size of the IR, be it in the separated or mixed condition. Our results suggest that there are no multiple sizes of second-order texture filters tuned to the different size of the texture figure, but instead a single filter of a fixed size for all figure sizes tested. Additionally, the integration size is on average larger for the detection task (4.1) than the segmentation task (3.3) for both participants SKS and DK. This is consistent with the modeling results we found in Experiments 1 & 2, where IR is larger for the detection task. A comparison was not made for participant SA, who was unable to produce physically meaning thresholds for the segmentation task to calculate the IR.


Grouped Detection
Grouped Segmentation
Mixed Detection
Mixed Segmentation
Integration Region
6
5
4
3
2
1
0
DK
12×15
16×20
20×25
Integration Region
6
5
4
3
2
1
0
SKS
12×15
16×20
20×25
Figure Size

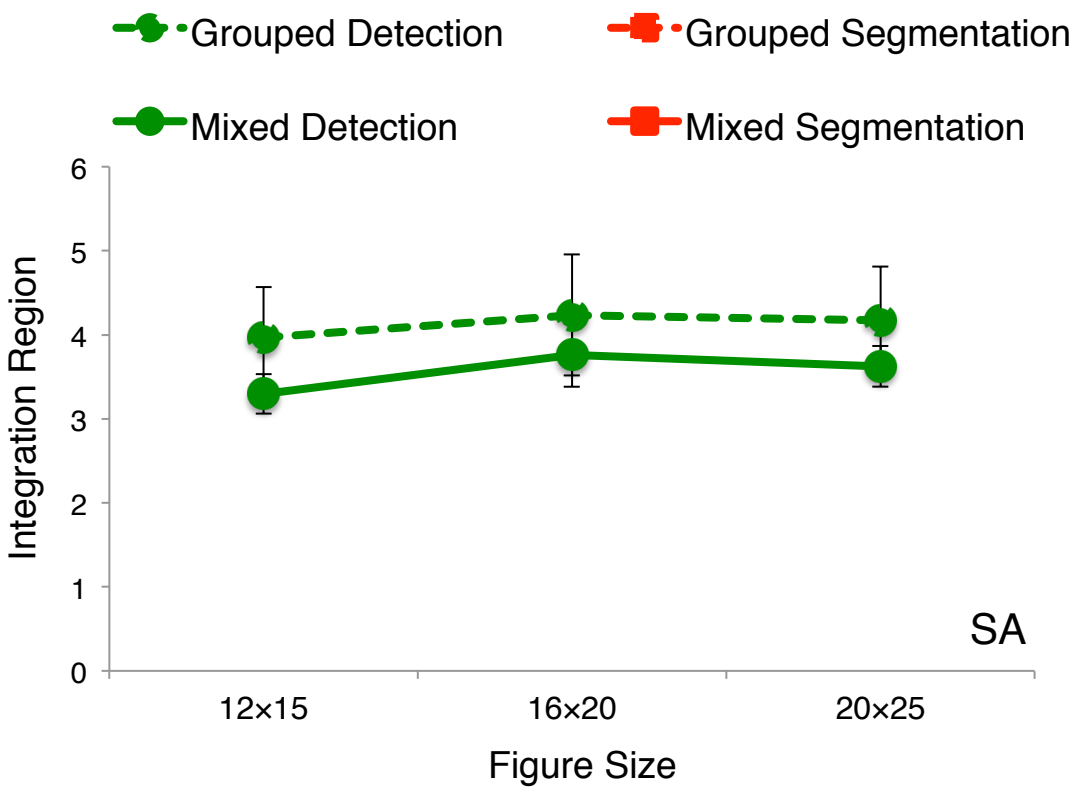


Figure 5.5. IRs as a function of figure size for the Detection (green circles) and Segmentation (red squares) tasks for the separated (dashed line) and mixed condition (solid line). For conditions in which thresholds were not available for all three profiles, the IR could not be computed. Error bars are the standard deviation, which were obtained by bootstrapping the data.

5.2 Experiment 4: Effect of Figure Aspect Ratio

As in Experiment 3 above, participants performed both the detection and segmentation tasks using stimuli with the three different orientation profiles. In this experiment however, we changed the aspect ratio of the figure size. We used the standard figure of 12×15 elements (1.25), and also figures of 9×15 (1.67), and 7×15 (2.14) elements. Thus, the height of the texture figures was maintained, while the width was decreased to produce figures with different aspect ratios. This experiment will allow us to investigate whether an increase in aspect ratio will result in smaller second-order texture filters.

We also investigated the ability of the visual system to adjust accordingly to figures with the different aspect ratios. That is, we ran participants in blocks where the figure aspect ratios were separated (one block will contain only one figure aspect ratio) or mixed (one block will contain all figure aspect ratios). If we find that the detection and segmentation of texture figures involve multiple second-order filters tuned to the different figure aspect ratios for the separated condition, but only a single fixed size second-order filter for the mixed condition, it would imply that the visual system was able to adapt to the aspect ratio of the figure being presented.

5.2.1 Methods

5.2.1.1. Participants

Three observers (DK, SKS and SA) participated in the experiment. DK and SKS were experienced psychophysical observers, while SA was a naive observer. All participants had corrected-to-normal vision.

5.2.1.2. Stimuli

The stimuli were created as described in Chapter 3, Section 3.3. However, in this experiment, we manipulated the aspect ratio of the figure size. We used the figure with 12×15 (3.0°×3.8°) elements [standard condition], and also included figures with 9×15 (2.3°×3.8°), and 7×15 (1.8°×3.8°) elements (see Figure 5.6). The aspect ratios of the figures were 1.25, 1.67, and 2.14 respectively.

For the Blur and Cornsweet stimuli, while the equation that determines the mean orientation of the line elements remains the same as described in Section 3.3, the number of lines over which the orientation change occurs is only 3 on either side of the border (see Figure 5.7).

This is in contrast to the standard setup where it is 5 (Cornsweet) and 6 (Blur) lines on either side of the border.

The reason we did this was because for the highest figure aspect ratio, the figure is narrower (only 7 lines wide), and the orientation change can occur over a maximum of 3 lines on either side of the border. For the purpose of consistency of orientation change between the three figure aspect ratios, we therefore had to reduce the number of lines over which the orientation change occurred.

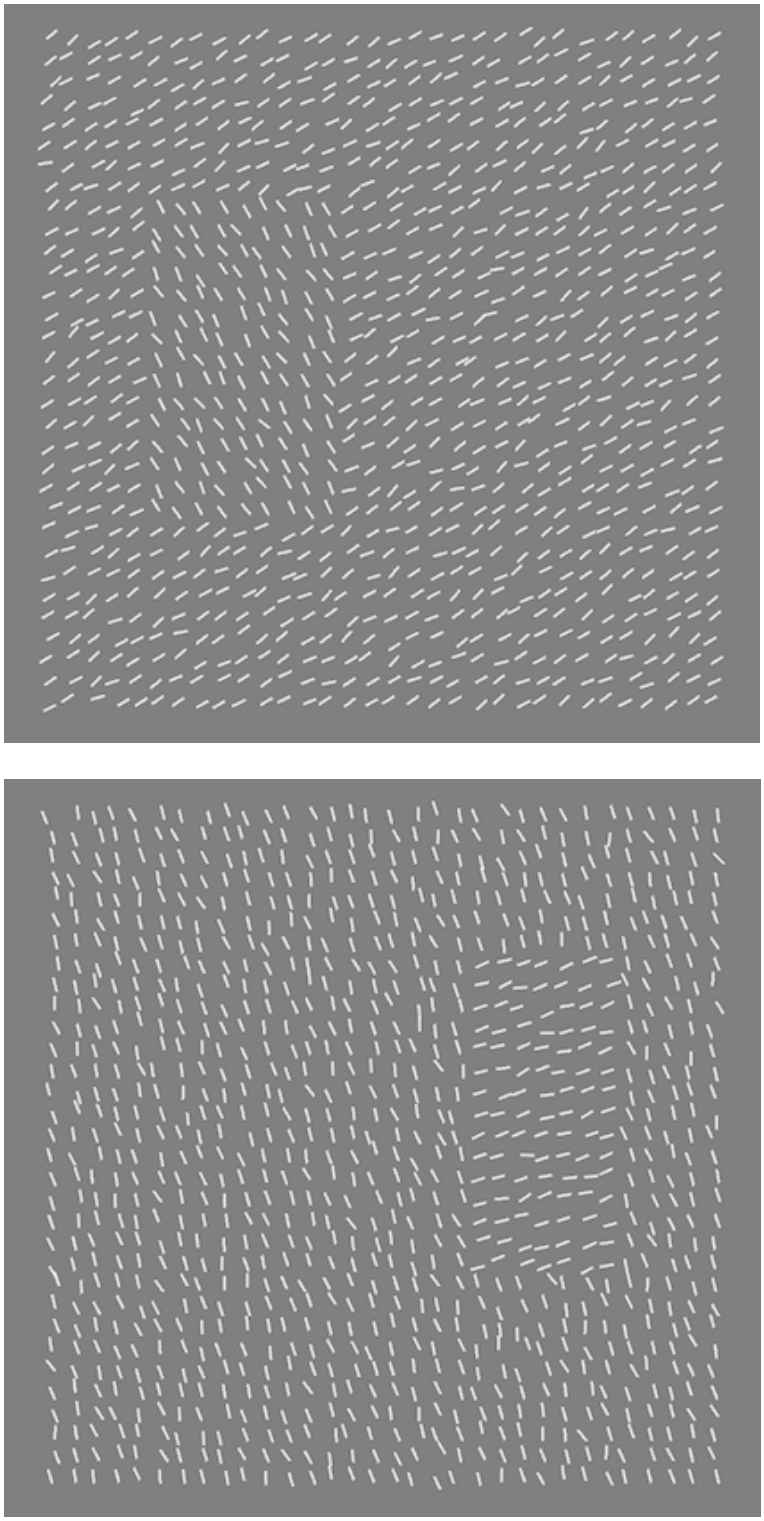


Figure 5.6. Stimuli with the 1.67 (top) and 2.14 (bottom) figure aspect ratio. Both stimuli are of the Block profile with 90° orientation contrast and 7.5° Standard Deviation for orientation jitter.

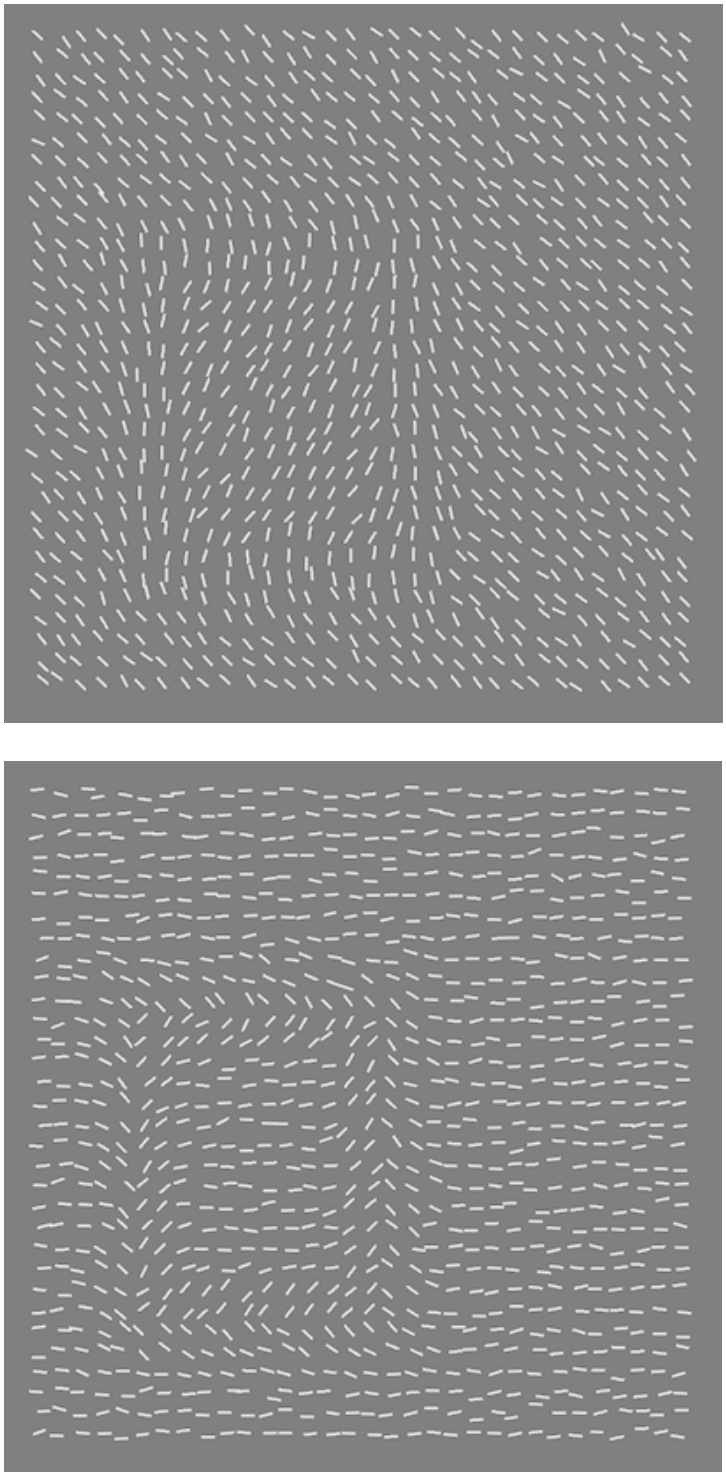


Figure 5.7. Stimuli with the 1.25 figure aspect ratio (12×15) for the Blur (top) and Cornsweet (bottom) profiles. Both stimuli show figures with 90° orientation contrast and 7.5° Standard Deviation for orientation jitter.

5.2.1.3. Procedure

As in Experiment 3, participants were required to perform the experiment twice – 1) with the different aspect ratios mixed within each block (mixed condition) and, 2) with only one aspect ratio within each block (separated condition). Participants performed both the Segmentation and Detection tasks (see Section 3.4), and all three orientation profiles (see Section 3.3) were used in the study.

The method of constant stimuli was used (see Section 3.1.1). For the mixed condition, in a given block of 180 trials, all orientation profiles and figure aspect ratios were presented in random order. The different tasks were presented in different blocks. For the separated condition, there were 150 trials per block, with the different orientation profiles presented in random order. The different tasks and figure aspect ratios were presented in different blocks. Each participant ran a total of 44 blocks over multiple days.

5.2.2 Results

Figures 5.8 and 5.9 show the orientation contrast thresholds of individual participants for the separated and mixed conditions respectively, with Figure 5.10 showing the average thresholds across participants. Once again, thresholds do not appear to be influenced by whether or not the figure aspect ratios were separated or mixed together, with the average difference in threshold being only 4.7%. This implies that the visual system does not over time adapt to the aspect ratio of the figure being presented. Additionally, only for the Block and Cornsweet profiles do the thresholds remain the same with the different figure aspect ratios tested for both the tasks. The Blur profile on the other hand, shows a decrease in orientation contrast thresholds when aspect ratio increases, but mostly in the segmentation task. That is, for the detection task, average thresholds of the Blur profile for the 2.14 figure

aspect ratio were 13.1% and 9.5% lower than the 1.67 and 1.25 figure aspect ratio respectively. However, for the segmentation task, average thresholds of the Blur profile for the 2.14 figure aspect ratio were 24.4% and 19.8% lower than the 1.67 and 1.25 figure aspect ratio respectively. For the Block and Cornsweet profiles on the other hand, differences in thresholds for the various figure aspect ratios ranged between 0.5% and 9.3%, which is considerable lower compared to the differences observed for the Blur profile of the Segmentation task.

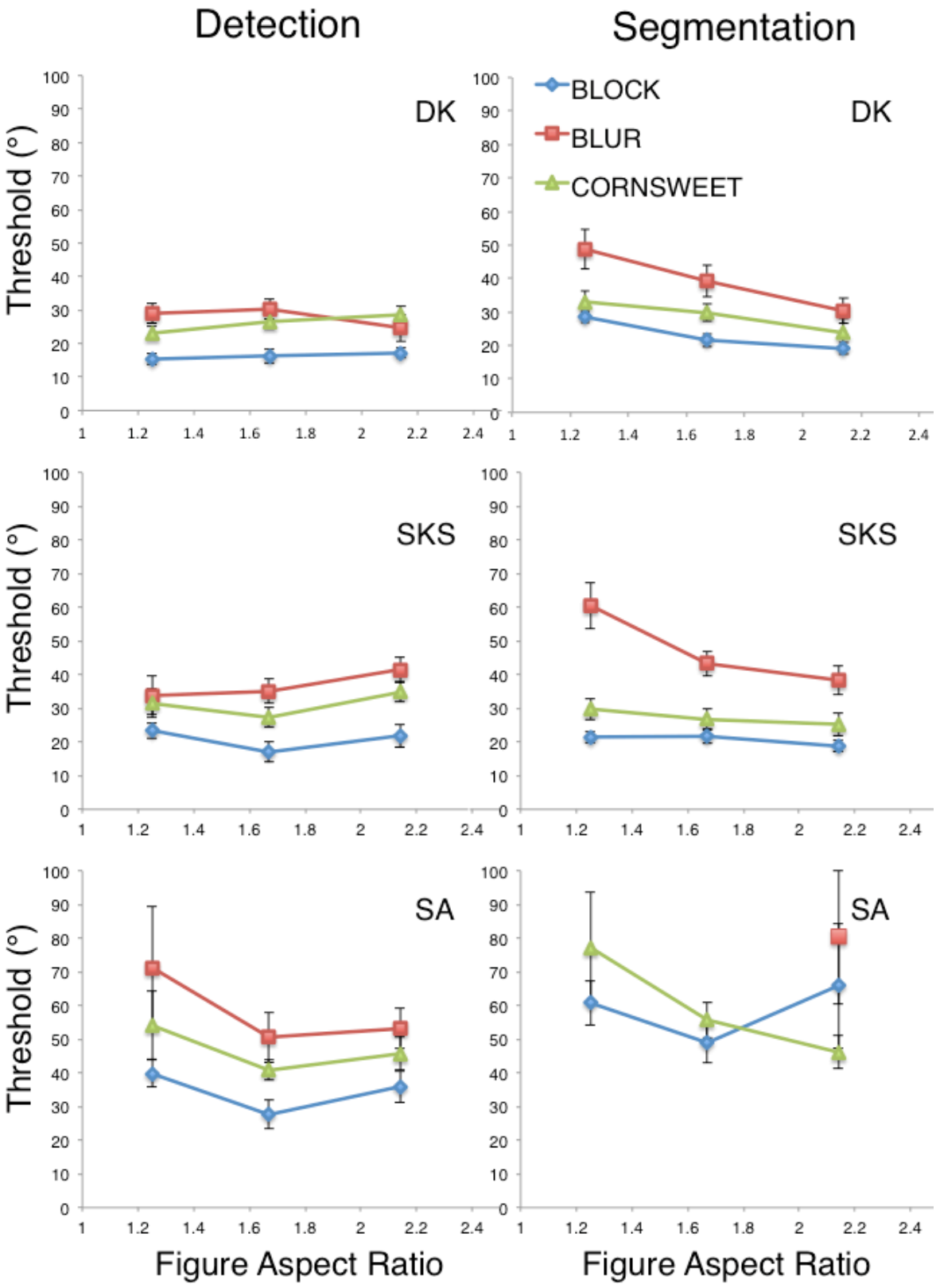


Figure 5.8. Thresholds as a function of figure aspect ratio for the Detection (left) and Segmentation (right) tasks for the separated condition. Error bars represent the 67% confidence interval.

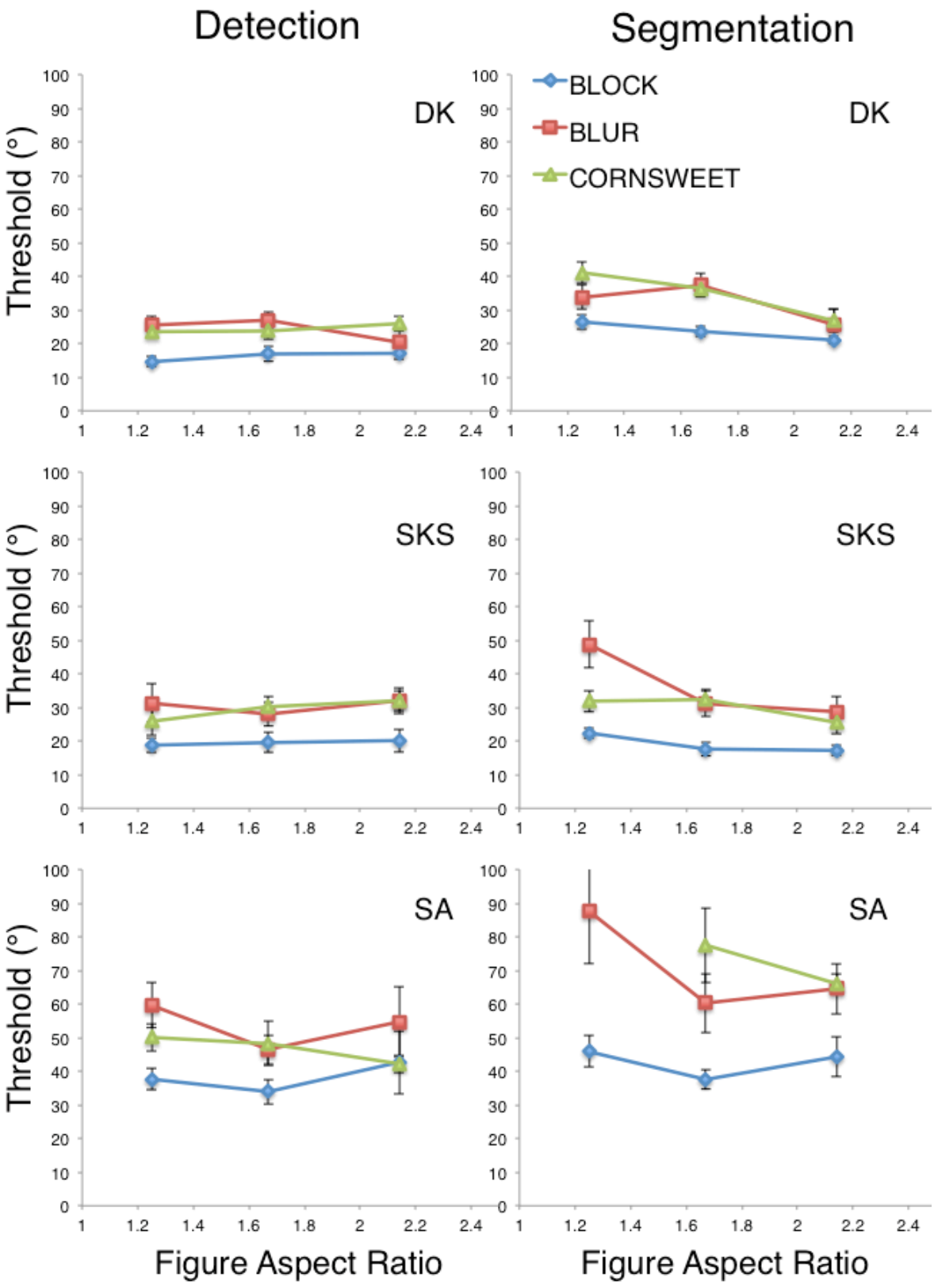


Figure 5.9. Thresholds as a function of figure aspect ratio for the Detection (left) and Segmentation (right) tasks for the mixed condition. Error bars represent the 67% confidence interval.

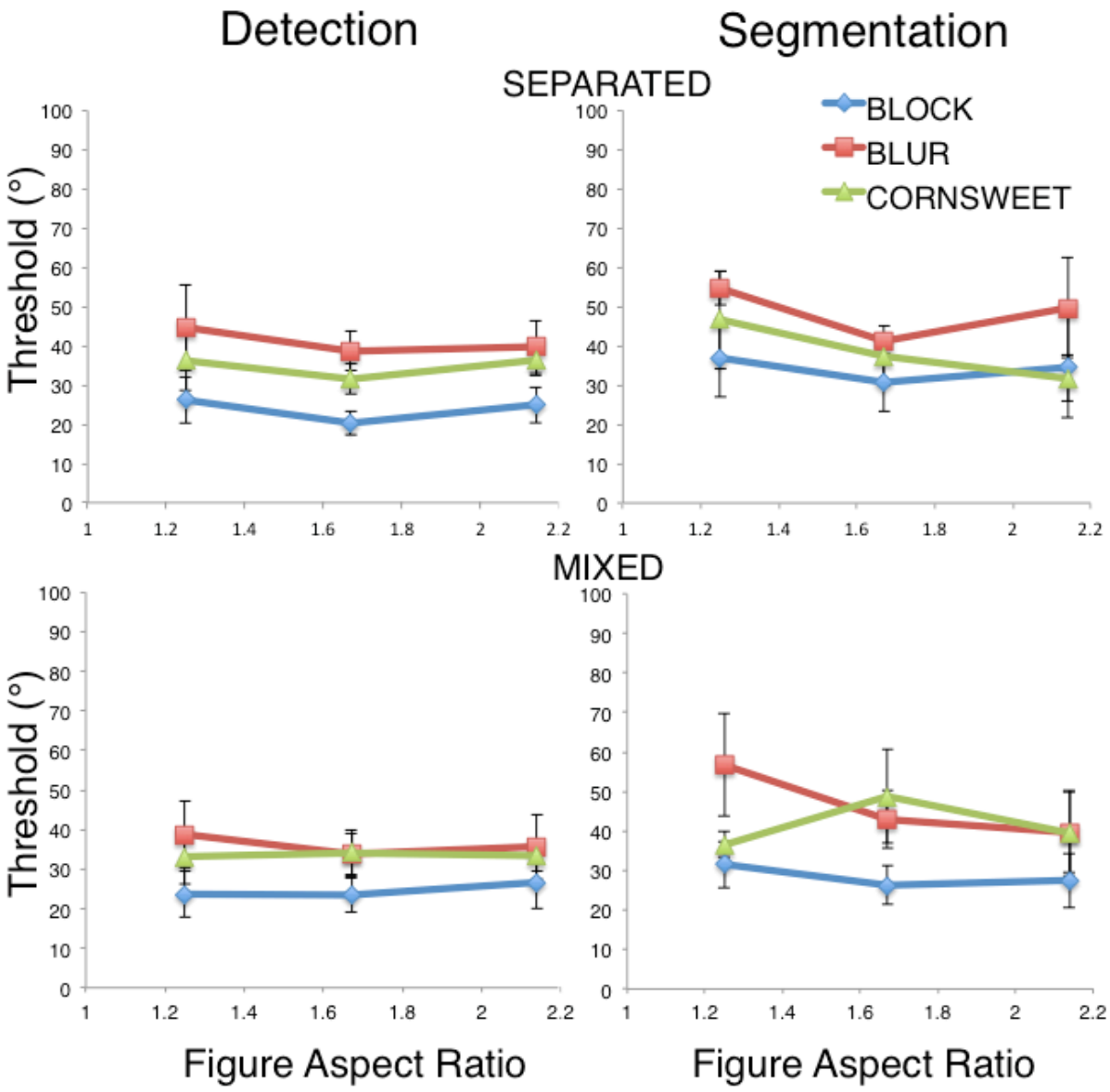


Figure 5.10. Average thresholds of all participants (*n*=3) as a function of figure aspect ratio for the Detection (left) and Segmentation (right) tasks. The different conditions are plotted on different rows. Error bars represent the standard error of the mean.

5.2.3 Modeling

The modeling of the Integration Region (IR) was done in the same manner as described in Section 3.5. However, for the 1.67, and 2.14 figure aspect ratios, the figure widths were smaller (7 & 9 columns) than the 1.25 figure aspect ratio (12 columns). As such, when modeling the results for the 1.67, and 2.14 figure aspect ratios, the orientation thresholds for

the three profiles could only be re-expressed in terms of the different mean orientations for raster units up to N=4 (Figure 5.11), instead of N=6.

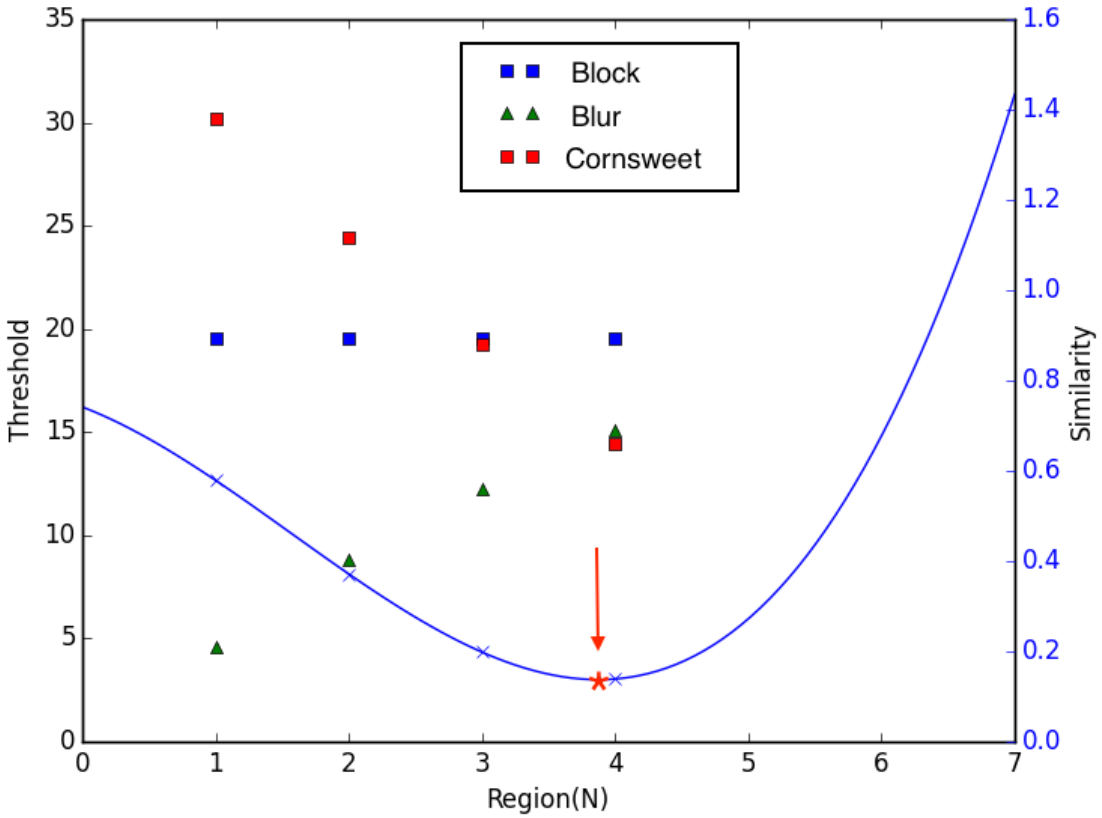
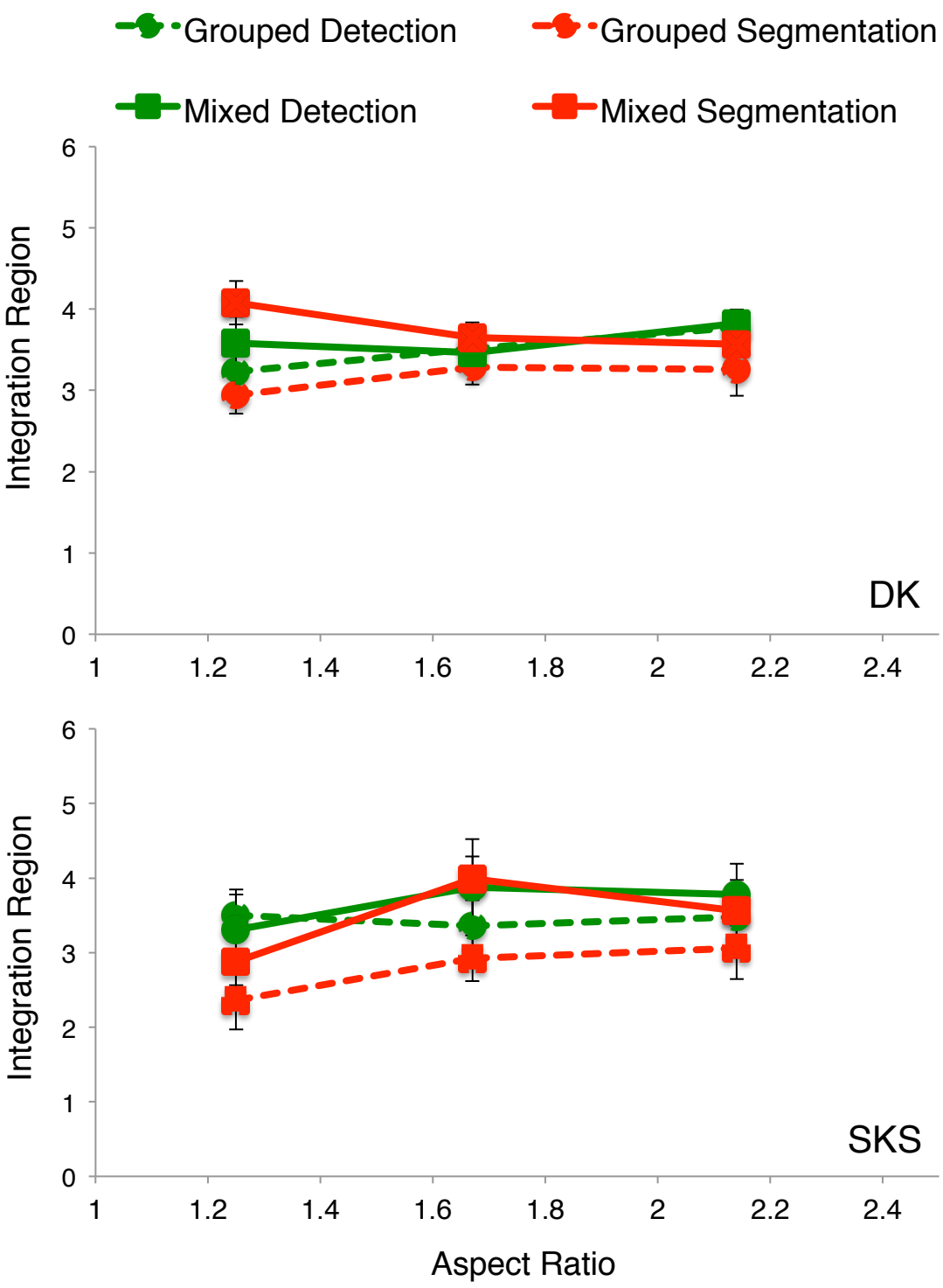


Figure 5.11. Example of the modeling for region sizes up to N=4. The IR (red *) is the point at which the thresholds for the 3 orientation profiles are most similar. Other than re-expressing the thresholds for region sizes up to a smaller raster size (4), no other changes were made to the modeling procedure.

The IRs are calculated for all conditions (2 tasks: segmentation and detection; 2 conditions: separated and mixed; and 3 aspect ratios: 1.25, 1.67, and 2.14) and plotted in Figure 5.12.

The results show that the IR does not change with the different figure aspect ratios in either the separated or mixed conditions. This implies that a single fixed size second-order texture filter serves to detect and segment texture figures with the different aspect ratios tested. Also, the IR on average for the detection task (3.5) is very similar to the segmentation task (3.3).

This is in contrast to the previous experiments (1 – 3), where the IR size is typically 1 raster unit larger for the detection task.



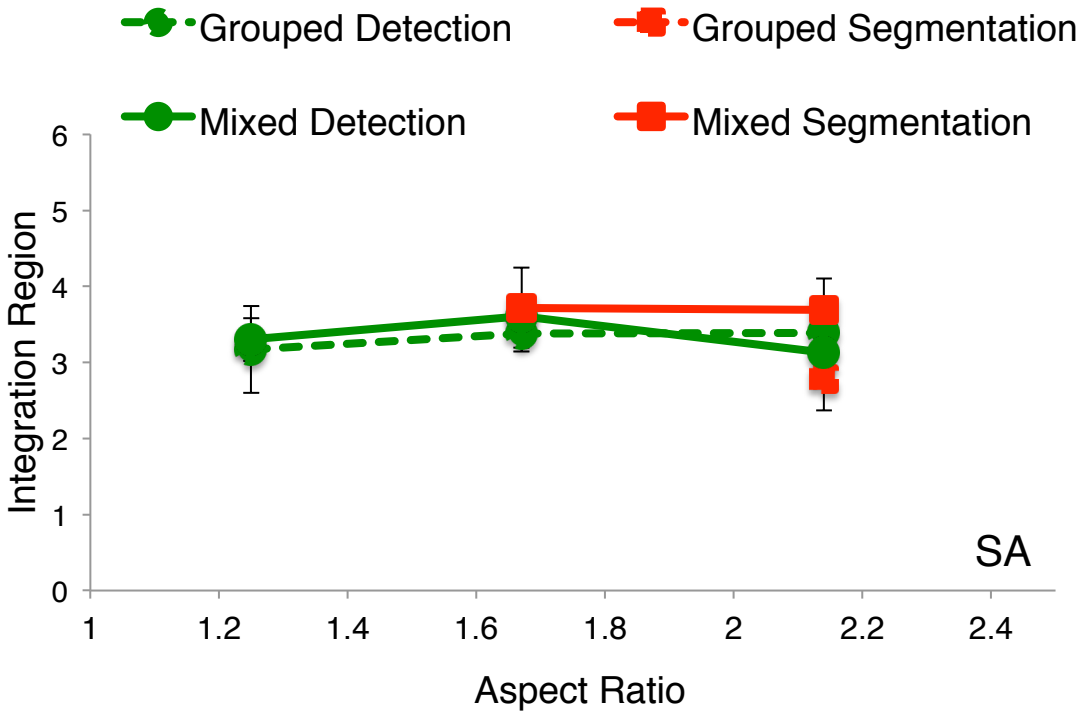


Figure 5.12. IRs as a function of figure aspect ratio for the Detection (green circles) and Segmentation (red squares) tasks for the separated (dashed line) and mixed condition (solid line). For conditions in which thresholds were not available for all three profiles, the IR could not be computed. Error bars are the standard deviation, which were obtained by bootstrapping the data.

5.3 Discussion

The purpose of the experiments presented in this chapter was to determine how changing the size and shape of a texture figure might affect performance in texture detection and segmentation tasks. The results of the study show the figure size and figure aspect ratio do not have much influence on the orientation contrast thresholds. The exception for this is the thresholds required to segment the Blur profile, where an increase in figure size and aspect ratio results in lower thresholds. Additionally, orientation contrast thresholds were not influenced by whether the different figure sizes and aspect ratios were mixed together or separated. We also found that the IR did not change with the different figure sizes and aspect

ratios tested, nor was it influenced by whether the figure types were separated or mixed within blocks. Finally, IR did vary between the detection and segmentation tasks, but only for the experiment in which we varied the figure size, not the experiment where we varied the figure aspect ratio.

As we discussed in Chapter 4, segmentation requires the border of the figure to be explicitly localised in order to perform the task. In contrast, detection merely requires an observer to be able to distinguish whether two regions appear different. We suggested that this distinction between the two tasks is the reason why performance for the Blur profile is severely impaired in the segmentation task. That is, orientation contrast thresholds increase on average by 21.6% and 16.9% respectively for the Block and Cornsweet profile, but it increases by 62.8% for the Blur profile. A possible explanation is that the Blur profile is characterised by its smooth texture variation i.e. no abrupt texture edge, which therefore makes localising the figure border more difficult compared to detecting a region that is different. Thus performance in the segmentation task of the Blur profile is markedly impaired compared to the other orientation profiles.

The results of Experiments 3 and 4 likewise show that compared to the detection task, thresholds for the segmentation of the Blur profile are much higher for the smaller figure size (12×15) and aspect ratio (1.25). However, as the figure size and figure aspect ratio increases, performance improves slightly. It implies that the increased size and aspect ratio of the figure aids in localising the border of the figure. It is plausible that as the aspect ratio of the figure increases (width becomes smaller), the shape of figure appears less square, which makes it easier to determine if the figure is oriented horizontally or vertically. This therefore results in the segmentation of the Blur profile improving with larger figure aspect ratios. Besides that, as the figure size increases, the centre region of the Blur figure (where the orientation

contrast in comparison to the ground orientation is the most) becomes larger, which would aid in segmenting the figure.

Moreover, to investigate whether there is a range of second-order texture filters tuned to different sizes or shape of a texture figure, we modelled the data in terms of an IR, which represents the size of the second-order filters. Our results were that the IR does not change with an increase in figure size or aspect ratio. This implies that unlike the range of second-stage filters that are tuned to different luminance spatial frequencies (Sutter, Sperling, & Chubb, 1995; Kingdom & Keeble, 1999), there is only a single fixed size second-order filter that serves the texture processing of different figure sizes and aspect ratio.

We also conducted the experiment with blocks where the different figure sizes and aspect ratios were separated (one block per figure size/aspect ratio) or mixed (all figure sizes/aspect ratios within one block). The rationale was that if the figure sizes were separated into individual blocks, our visual system would adapt over time to the figure size/aspect ratio being presented. This theoretically could have resulted in the second-order texture filter getting tuned to the appropriate size to best process the texture figures. However, even in the separated condition we found that the IR is not affected by the figure size or aspect ratio.

We do find that the IR of detection task is on average larger than in the segmentation task, but only for Experiment 3 [detection=4.1; segmentation=3.3] and not Experiment 4 [detection=3.5; segmentation=3.3]. We previously proposed in Chapter 4 that a larger IR would better serve the processing of texture stimuli for a detection task, where sampling a larger portion of the stimulus would be more efficient in detecting the presence of two distinct regions to the texture. However, a smaller IR would be most efficient in explicitly forming an edge boundary to localise the figure. On average, the IR for detection task is also larger than the segmentation task for Experiments 1 [detection=3.6; segmentation=2.6] and 2

[detection=3.3; segmentation=2.3], though they are smaller than the IRs found in Experiment 3. A possibility for the overall larger IRs for Experiment 3 is the size of the texture grid, which was 40×40 instead of 32×32 like in the previous experiments. To quantify this, the 40×40 grid is 1.25 times larger than the 32×32 grid. Therefore, if the IR of Experiment 1 were also 1.25 times larger, the IR should be 3.3 and 2.6 for the detection and segmentation task respectively, which is very close to the IR of Experiment 3.

A key point to note about these experiments is that we have not been able to decouple other effects that are associated with the change in figure sizes and aspect ratios. For example, as we increased the size of the figure, the number of elements within the figure also increased. Likewise, as we increased the figure aspect ratio, both the size and number of elements of the figure decreased. However, Nothdurft (1985a) found that pattern (numbers 0–9) discrimination was not affected by the overall size of the pattern, nor by the number of elements, though line spacing did influence the discrimination of the patterns (compare curves I, II, and III of Figure 7a in Nothdurft 1985a). Therefore, it is possible that the confound of figure size and number of elements would not affect the results of our experiments.

Chapter 6

Effects of Aging on Spatial Integration in Texture Segmentation

In this chapter, we are interested in investigating whether the processing of texture stimuli declines with old age. In Section 6.2, we specifically question how the aging visual system responds to the spatial frequency content during texture segmentation. This idea of studying how the elderly population performs when manipulating the spatial frequency in an orientation-based texture task stems from four reasons. First, studies have found that while older adults show decreased contrast sensitivity to medium and high spatial frequency gratings, low spatial frequency gratings are not affected by age (Owsley et al., 1983). Second, consistent with the view that impairment in visual performance of the elderly is partially due to changes in the physiological properties of the visual neurons, Schmolesky et al., (2000) found evidence of decreased sensitivity to orientation selective cells. Third, in the vision and aging literature, there have been somewhat contradictory findings about the performance of older adults. This is because some studies show that older adults outperform young adults with large stimuli (Betts et al., 2005), while others show that they outperform them with smaller stimuli (Hutchinson et al., 2014). Therefore, our model, which measures the size of a region in which information is extracted form, can perhaps add to this debate whilst also adding to our aim of investigating whether there are a range of second-order filters tuned to different properties (also see Chapter 5). And finally, to the best of our knowledge, no research has investigated the effects of aging in an orientation-based texture segmentation task. The closest are studies that investigate orientation discrimination in young and old adults (Delahunt et al., 2008; Govenlock et al., 2009; Casco et al 2017; Wang et al., 2018),

and Kurylo (2006) who investigated the feature of orientation in perceptual organisation. This will be discussed further in Section 6.2.

However, before we conducted the study that investigates the effect of aging on orientation-based texture segmentation, we first piloted the study on a few participants at a range of different luminance spatial frequencies. This was done because studies on texture perception have demonstrated the importance of spatial frequency content in the segregation of texture regions. Landy and Bergen (1991) for instance, have shown that performance in a shape discrimination task improves with higher spatial frequencies. However, their study only tested 3 spatial frequencies – 0.75, 1.5, and 3.0cpd, which leaves out a large range of higher spatial frequencies. Besides that, the human visual system's contrast sensitivity function is characterised as having varying degrees of sensitivity to different spatial frequencies. That is, if our visual system is good at detecting a particular spatial frequency, the contrast needed to see the grating will be low. On the other hand, if our visual system is poor at detecting a particular spatial frequency, the degree of contrast needed to see the grating will be higher. Thus, the issue of visibility potentially confounds the results. This was therefore explored first in the study presented in section 6.1.

6.1 Pilot Work with Gabor Elements with Varying Spatial Frequencies

For the sake of simplicity, and since it is a pilot study, we only used one of the three orientation profiles – Block profile, and manipulated the spatial frequency of the Gabor elements (between 2.5 and 18.9 cpd). As mentioned above, visibility of the gratings could have an influence on performance. To control for this, we measured each individual participants' contrast sensitivity to each of the spatial frequencies used by having them perform an Orientation Discrimination task. According to Pointer and Hess (1989), contrast sensitivity is affected by the eccentricity of the stimuli. As such, when participants performed

the Orientation Discrimination task, the stimuli appeared away from the fovea (1/2 the distance of the texture stimulus from fixation). The participants then completed a segmentation task twice: 1) with the contrast of the Gabor elements at maximum (Maximum Contrast), and 2) with the contrast of the Gabor elements at a fixed number of decibels above each individuals' contrast thresholds (Matched Contrast). The purpose of the Matched Contrast experiment was to equate the visibility of the Gabor elements that comprise the texture grid. If we see a difference in performance between the Maximum Contrast and Matched Contrast, it would imply that the visibility of the Gabor elements was also playing a role in segmentation performance.

6.1.1 Methods

6.1.1.1 Participants

Three observers (DK, SKS, and NA) participated in the experiment. DK (age = 55) and SKS (age = 26) were experienced psychophysical observers, while NA (age = 21) was an intern who was a naïve observer. All participants had corrected-to-normal vision.

6.1.1.2 Display

The display was the same as described in Section 3.2, with the exception that the pixel resolution was 0.76 arcminute per pixel when viewed from a distance of 175cm.

6.1.1.3 Stimuli

The stimuli for the Maximum Contrast and Matched Contrast study comprised Gabor elements on a grey background (see Figure 6.1). Each stimulus was an array of 9.1°×9.1° containing 256 Gabor elements arranged in a 16×16 grid. As in the previous studies, the figure of the texture stimulus was the region in which the orientation of the Gabor elements

was different from the background. For this pilot study, only the Block profile was used, which was of a 6×8 figure grid spanning 3.4°×4.6°. The peak luminance spatial frequencies used were 2.53, 3.79, 5.68, 8.37, 10.58, 12.63, and 18.98cpd. The contrast of the Gabor elements will be described in detail in the procedure below.

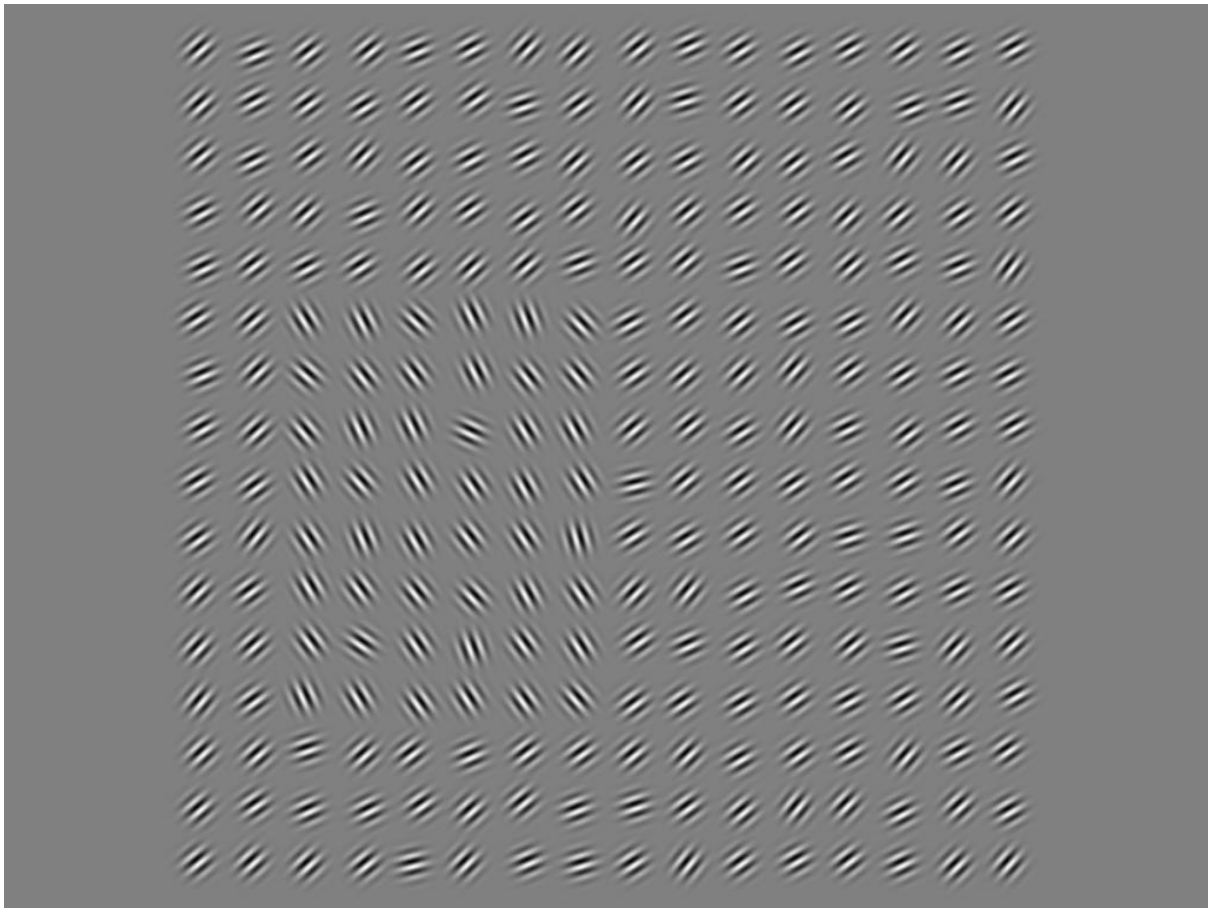


Figure 6.1. Example of the texture stimuli used in the Segmentation Task. The figure is of a Block profile with 90 degrees orientation contrast. The peak luminance spatial frequency is 8.37 cpd, and the contrast was at 1 (maximum contrast).

The border of the figure region was set to appear at least 2 lines away from the edge of the stimulus array i.e. figure appears on the 3rd line onwards. Additionally, position and orientation jitter was added by drawing samples from a Gaussian distribution with standard deviations of 2 pixels (position) and 7.5° (orientation). Where Gabor elements overlapped, the luminance values were averaged.

Participants also performed an Orientation Discrimination task to measure their contrast sensitivity. For this task, only one Gabor element was presented on a circle (see Figure 6.2). The Gaussian standard deviation (σ) of the Gabor element was 0.11° (same as the Gabor elements for the Maximum Contrast and Matched Contrast study), and the distance between the fixation cross and the Gabor element was 2.3°. This distance is equivalent to a quarter of the width/height of the texture stimulus, or half the width/height if viewed from the central fixation. The same luminance spatial frequencies as above were used.

Figure 6.2. Example of the single Gabor element as seen in the Orientation Discrimination task. The contrast of the Gabor element seen here is at 1 (maximum contrast), and the peak luminance spatial frequency is at 5.68 cpd.

6.1.1.4 Procedure

The experiment was broken up into three separate parts: Orientation Discrimination task, Maximum Contrast study, and Matched Contrast study. The Orientation Discrimination task was used to measure contrast sensitivities for the different luminance spatial frequencies tested. A two-alternate forced-choice (2AFC) procedure was employed, whereby the orientation of the Gabor element could either be vertical (0°) or horizontal (90°), and the participant had to discriminate the orientation of the element (see Figure 6.2). On each trial, the fixation cross was displayed for 550ms, followed by the Gabor element for 108ms. The method of constant stimuli was used (see section 3.1.1), with the seven log-spaced levels of intensities being the contrast of the Gabor element. In a given block (10 in total) of 196 trials, all luminance spatial frequencies were tested in random order, and contrast thresholds for each spatial frequency were generated from 240 trials. Feedback in the form an auditory tone was given for incorrect responses.

The task used for the Maximum Contrast and Matched Contrast study was the segmentation task (2AFC procedure, see section 3.4.2), which measures the threshold of orientation contrast. The method of constant stimuli was used to determine thresholds (see section 3.1.1). During any given block of 280 trials, all luminance spatial frequencies were presented in random order.

The only difference between the Maximum Contrast and Matched Contrast study was the contrast value. For the Maximum Contrast, the contrast was at its maximum (1) for all luminance spatial frequencies. For the Matched Contrast study, new contrast values were used, which differed between the different luminance spatial frequencies. These new contrast values were obtained by increasing the contrast (derived from the Orientation Discrimination task, see Figure 6.3) by fixed decibels.

6.1.2 Results

6.1.2.1 Orientation Discrimination of a Single Gabor Element

Contrast threshold (solid line) for the Gabor element is plotted as a function of luminance spatial frequency in Figure 6.3 below. The new contrast values (× in Figure 6.3) were obtained by increasing the contrast by 20dB (SKS), 15dB (DK), and 10dB (NA). The formula for calculating contrast in decibels were:

$$\mathrm{dB} = -20 \log_{10}(\mathrm{threshold})$$

The reason for the different increases in decibels is so that the participants could be tested at a greater range of spatial frequencies. For instance, if DK and NA were to use 20dB increase in contrast, they would not be able to be tested at the 10.6cpd grating, as the new contrast value would be more than 1 (not physically possible). Although the amount of increase in contrast differs between the participants, the new contrast values are still **above** threshold for all participants, thus still allowing us to investigate if the visibility of the Gabor elements was influencing texture segmentation. Participant SKS was not able to produce a physically meaningful threshold for the 18.9cpd grating, while participants DK and NA were not able to produce physically meaningful thresholds for the 12.6 and 18.9cpd gratings. Participants DK and NA produced the inverted u-shaped curves typically seen in contrast sensitivity functions, but participant SKS did not.

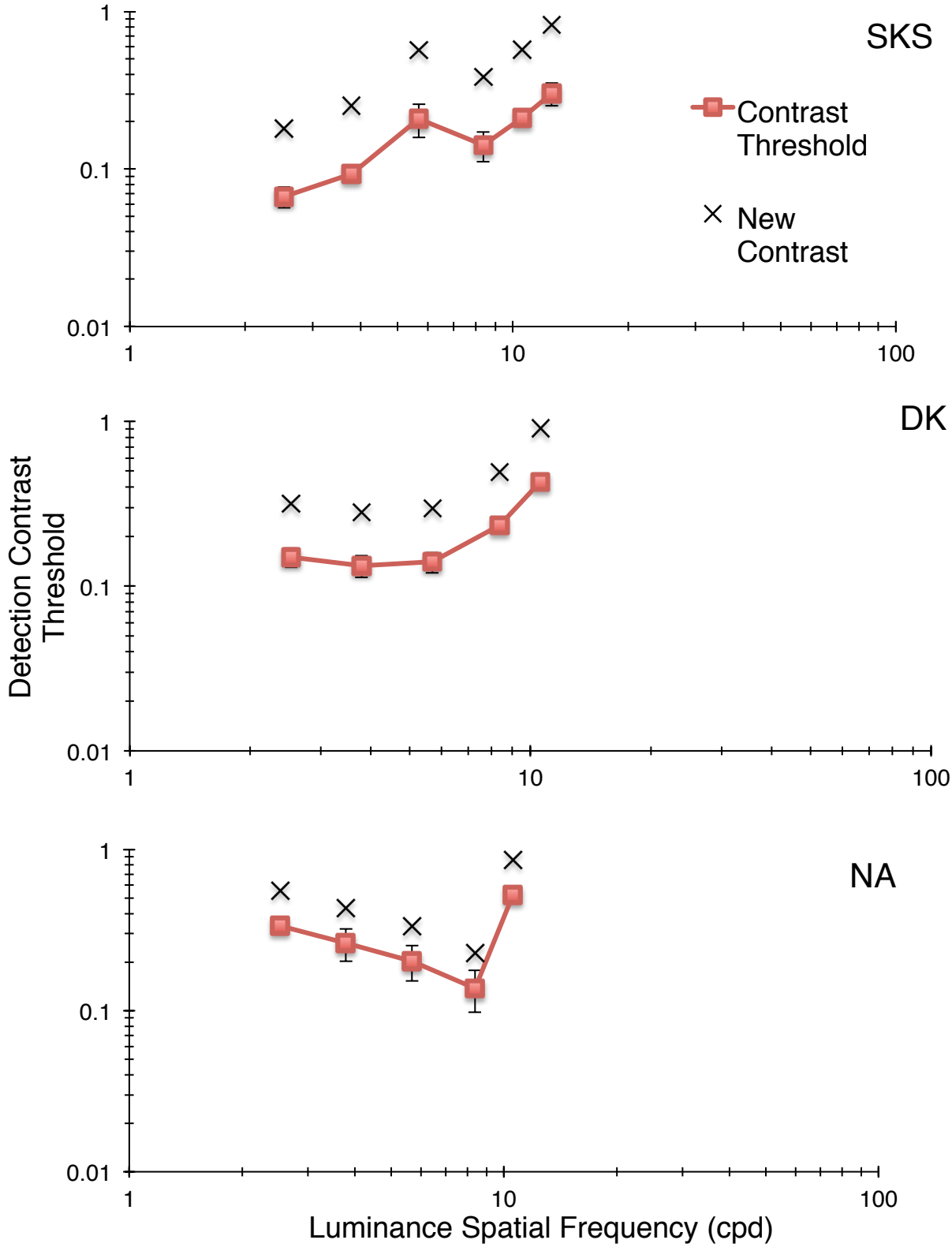


Figure 6.3. Contrast thresholds (solid line) for the discrimination of a single Gabor element. The × represents the contrast of the texture stimuli used in the Segmentation Task for the Matched Contrast study. Error bars on contrast thresholds represent 67% confidence interval.

6.1.2.2 Maximum Contrast and Matched Contrast

Orientation thresholds were in terms of orientation contrast between figure and ground. Figure 6.4 shows the results for both the Maximum Contrast and Matched Contrast study for all participants.

The most prominent finding of these studies is that even when the contrast was matched equally (i.e. equating the visibility) for the Matched Contrast study, the orientation threshold values were very similar to the Maximum Contrast study, with a difference of only 4.2%. This is true for both participants SKS and DK at all luminance spatial frequencies tested, and for participant NA at the lower spatial frequencies tested. For participant NA, orientation contrast thresholds at spatial frequencies of 8.37 and 10.58cpd are on average 2.2× higher for the Matched Contrast task compared to the Maximum Contrast task. However the error bars for those spatial frequencies were also very large, which could account for the difference in thresholds. These results tell us that the change in performance for the different spatial frequencies in a texture segmentation task is unlikely to be due to changes in visibility. Apart from that, the orientation threshold shows a u-shaped curve for all participants, with lowest thresholds being generated for the luminance spatial frequency of 5.68cpd for all participants.

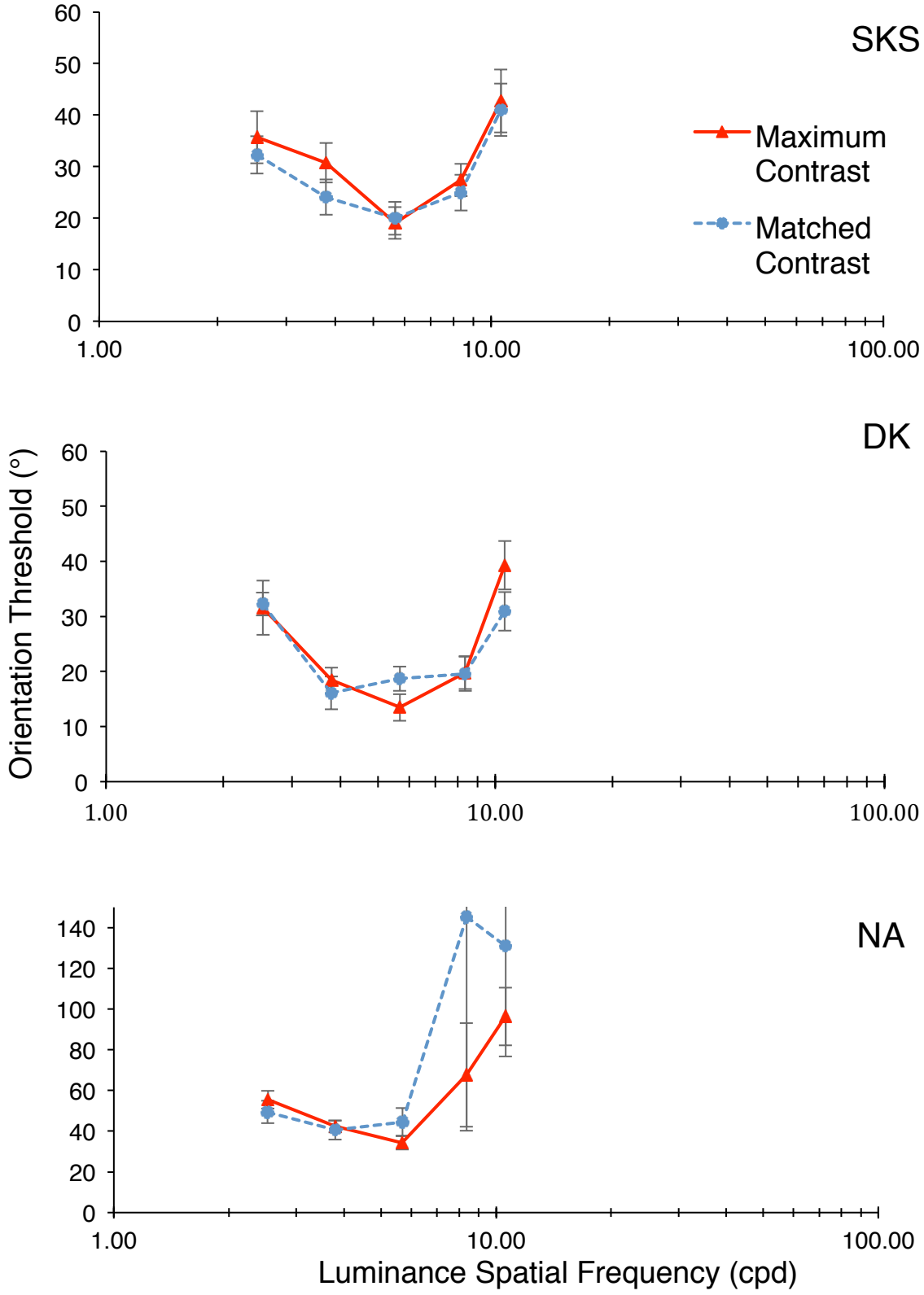


Figure 6.4. Orientation contrast thresholds for the segmentation of texture stimuli for the Maximum Contrast (solid line) and Matched Contrast (dashed line) study. Error bars represent the 67% confidence interval.

6.2 Experiment 5: Assessing Segmentation Performance of Young and Old Adults

In this study, we measured old and young participants' performance in a segmentation task. We used two spatial frequencies – 1cpd and 4cpd, as old and young participants' sensitivity to gratings at 1cpd are similar to each other, but differ at 4cpd (e.g. Owsley et al., 1983). We also measured participants' contrast thresholds in an orientation discrimination task for those spatial frequencies.

The findings of the effects of aging on vision are widely disparate, but there is a trend suggesting that some functions are spared throughout adulthood. Snowden and Kavanagh (2006) found that irrespective of spatial frequency, older participants were worse at detecting the direction of motion of a grating. This was in accordance with some studies (e.g. Habak & Faubert, 2000; Faubert 2002; Tang & Zhou, 2009; Bertone et al., 2011), where older adults perform worse overall compared to younger adults simply because of the second-order stimuli used. However, other studies (e.g. Allen et al., 2010; Hutchinson et al., 2014; Betts et al., 2005) have shown instances in which older people are actually better (i.e. while young participants show impaired performance, older participants do not) than younger people. In fact, Snowden and Kavanagh (2006) also show instances in which old people perform just as well as young people.

For our study, if older adults consistently perform worse, it could imply that they do in fact have trouble processing second-order (in our case texture-defined) stimuli. Alternatively, it is also possible that older adults are able to process second-order stimuli, and since older adults have shown an intact ability to process orientation information (e.g. Delahunt et al., 2008), performance will only be worse for stimuli in which there are loss of inputs of high spatial frequencies.

6.2.1 Methods

6.2.1.1 Participants

This experiment was conducted in Nottingham, UK with a population of young and old participants (ethics approval number: S938). We took into consideration the difficulties older adults may face while performing this experiment, and some changes were made to the experimental design to ease their burden. These will be highlighted in the following sections.

30 individuals, 15 old (M_{age}=69.9 years, *SD*=3.5) and 15 young (M_{age}=23.9 years, *SD*=1.4), participated in the experiment. All participants had normal or corrected-to-normal vision.

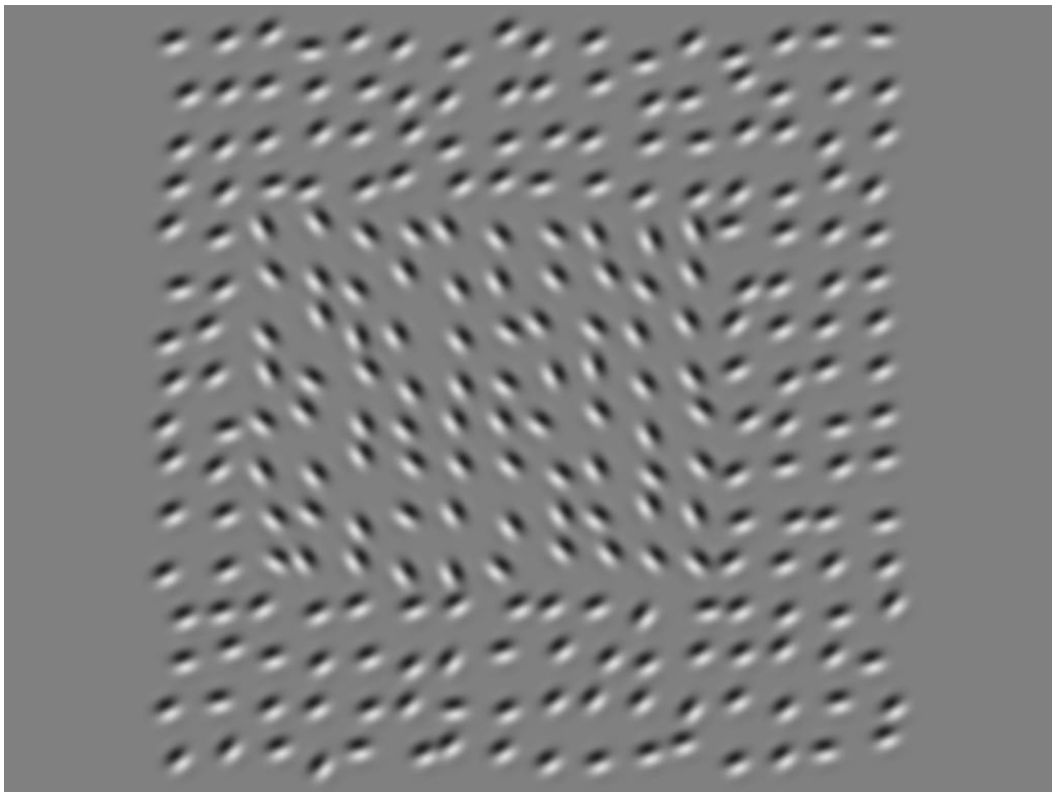
6.2.1.2 Display

The stimuli were generated using PsychoPy 1.83.04 (Pierce, 2007) on an Apple Mac Pro, and displayed on a ViewSonic Professional Series P225f CRT monitor that had been gamma corrected. The screen resolution was 1280×1024, and had a refresh rate of 85Hz. At a viewing distance of 54cm, the pixel resolution was 1.8 arcminutes per pixel.

6.2.1.3 Stimuli

The stimuli for the Segmentation task comprised Gabor elements on a grey background (see Figure 6.5). Each stimulus was an array of 21.6°×21.6° containing 256 elements arranged in a 16×16 grid. As in the previous studies, the figure of the texture stimulus was the region in which the orientation of the Gabor elements was different from the background. In this study, the Block, Blur, and Cornsweet profiles were used, with the 8×10 figure grid spanning 10.8°×13.5°. The equation used to produce the orientation of the Gabor elements for the Blur and Cornsweet profiles were similar to equation described in Section 3.3. However, in this study, k and f (parameters for the Blur and Cornsweet profile respectively) were 45 and

0.01309, and for equations 2 and 3 (Cornsweet profile), *s* was 45 pixels. Additionally, the number of lines over which the orientation change occurs is only 3 (Cornsweet) and 4 (Blur) on either side of the border. This is in contrast to the standard setup where it is 5 (Cornsweet) and 6 (Blur) lines on either side of the border. The purpose of this was to accommodate the smaller 8×10 figure grid. In this study, only two peak luminance spatial frequencies were used: 1.05cpd and 4.20 cpd. The contrasts of the Gabor elements were set at its maximum (1).



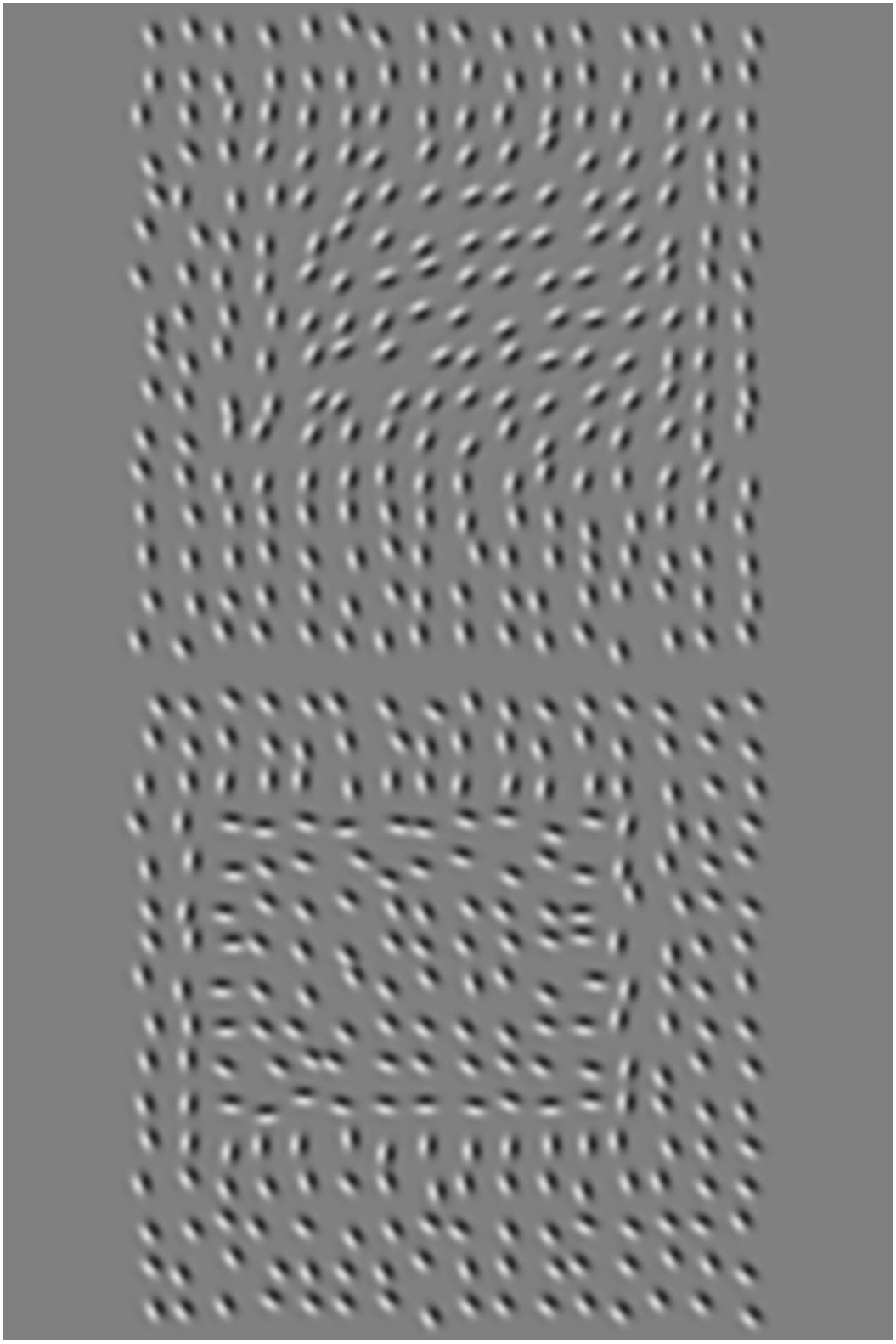


Figure 6.5. Example of the texture stimuli used in the Segmentation Task. The figure of the texture had three different profiles: Block (top), Blur (middle), and Cornsweet (bottom). The images shown here are example of stimuli with 1.05cpd peak luminance spatial frequency, and 80 degrees orientation contrast.

The border of the figure region was set to appear at least 2 lines away from the edge of the stimulus array i.e. figure appears on the 3rd line onwards. Position and orientation jitter was added by drawing samples from a Gaussian distribution with standard deviations of 6 pixels (position) and 7.5° (orientation). Where Gabor elements overlapped, the luminance values were averaged.

For the Orientation Discrimination task, only one Gabor element (Gaussian envelope with σ of 0.25°) was presented randomly on a circle that was 5.5 degrees away from the fixation cross (see Figure 6.6). As before, this distance is equivalent to a quarter of the width/height of the texture stimulus, or half the width/height if viewed from the central fixation. Two peak luminance spatial frequencies were used: 1.05cpd and 4.20 cpd.

Figure 6.6. Example of a single Gabor element as seen during the Orientation Discrimination task. The contrast of the Gabor element seen here is at 1 (maximum contrast), and the peak luminance spatial frequency is 4.20 cpd. Note how the fixation cross is dimmer than in Figure 6.2. The purpose of this was to help the older adults disengage attention from the fixation cross.

6.2.1.4 Procedure

The experiment was broken up into two parts: the Orientation Discrimination task, and the Segmentation task. The procedure of the Orientation Discrimination task was similar to the methodology described above (section 6.1.1.4). The only differences are: 1) the QUEST type (Watson & Pelli, 1983) interleaved staircase method was used, see Section 3.1.2, 2) in a given block of 240 trials, both luminance spatial frequencies were used, 3) contrast thresholds were generated from 120 trials each, 4) the sequence of trial begins with a fixation for 550ms, which then dims as the Gabor element appears for 1000ms, followed by a question mark appearing on-screen until a response is made, 5) the first 12 trials of the experiment were made especially easy i.e. high contrast, and these responses were not taken into account during the analysis.

For the segmentation task, the procedure was the same as described in the above study (Section 6.1.1.4). The differences between the previous studies and this study are: 1) the QUEST type (Watson & Pelli, 1983) interleaved staircase method was used, see Section 3.1.2, 2) in a given block (5 in total) of 168 trials, both luminance spatial frequencies and all three profiles were used, 3) orientation thresholds were generated from 120 trials each, 4) the sequence of trial begins with a fixation for 550ms, followed by the stimulus for 1000ms, a blank screen for 47ms, and a mask for 23ms, 5) the first 12 trials of the sequence were made especially easy i.e. orientation contrast >85°, and these responses were not taken into account during the analysis, 6) 14% of the total number of trials were easy to segment with high orientation contrast figures, which were also not taken into account during analysis, 7) participants performed a practice block (with difficulty of trials gradually increasing) before beginning the task.

Experiment 5 has some difference to the Pilot study, but these were a result of us trying to make the experiment more suited for the older adults. For instance, an experiment that is **longer** than 1–1.5 hours (which is the approximate duration of this study) would be taxing on the elderly participants. Thus, we choose to use the interleaved staircase method in this study, as it requires fewer trials to estimate a threshold. For the orientation discrimination task, we dimmed the fixation cross when the Gabor element was displayed to help the older adults disengage their attention from the fixation cross. Furthermore, we displayed a question mark when it was time to respond (i.e. after offset of Gabor element) to serve as a visual cue. For both tasks, the first few trials were made easy and discarded from analysis because older adults would typically make mistakes (like pressing the wrong keys) unrelated to their ability to perceive the stimuli. For the segmentation task, 14% of the trials were made easy so as not to demotivate the participants.

6.2.2 Results

6.2.2.1 Orientation Discrimination of a Single Gabor Element

Four participants from the older age group were not able to produce contrast thresholds that were physically meaningful (i.e. less than 1), and were removed from the statistical analyses (ANOVAs) for this task (Orientation Discrimination) and the Segmentation Task. Figure 6.7 shows the average contrast thresholds of the remaining participants for the Gabor element at 1.05cpd and 4.20 cpd (luminance spatial frequency).

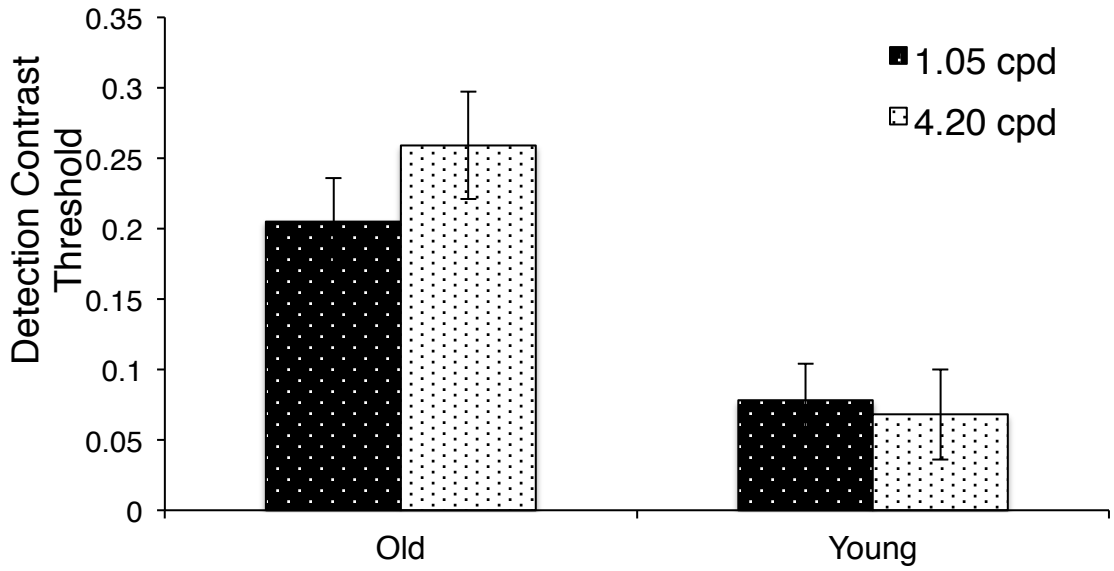


Figure 6.7. Average contrast thresholds of the older and younger participants for the Orientation Discrimination of a single Gabor element at 1.05cpd and 4.20cpd. Error bars represent the standard error of the mean.

A 2 (luminance spatial frequency: 1.05 and 4.20cpd) × 2 (age group: old and young adults) Mixed-Design ANOVA was used to analyse the results. An overall main effect of age (between subject) was found $F(1,24)=13.290$, $p=0.001$, showing that older participants had higher contrast thresholds compared to younger participants ($M_{old} = 0.232$; $M_{young} = 0.073$). An interaction between age and luminance spatial frequency was also found $F(1,24)=6.171$, $p=0.020$, where Bonferroni corrected pairwise comparisons showed that while luminance spatial frequency did not affect the contrast threshold of the younger participants ($M_{LSF}=0.078$; $M_{HSF}=0.068$, $p=0.547$), it did affect the contrast thresholds of the older participants ($M_{LSF}=0.205$; $M_{HSF}=0.259$, $p=0.011$). There was no main effect of luminance spatial frequency $F(1,24)=2.856$, $p=0.104$.

6.2.2.2 Segmentation of Texture Stimuli

Weibull functions were fitted to the psychometric functions, and thresholds at 70% correct level were found. The 70% correct level was used because at 82% correct, a majority of the

participants could not produce physically meaningful thresholds (i.e. below 90°). Despite this, the data of 7 older participants and 1 younger participant were still not useable. Orientation thresholds for the Segmentation task are shown in Figure 6.8, and Table 6.1 shows the findings of the 2 (luminance spatial frequency: 1.05 and 4.20cpd) × 2 (age group: old and young adults) × 3 (orientation profile: Block, Blur, and Cornsweet) Mixed-Design ANOVA conducted on the useable data.

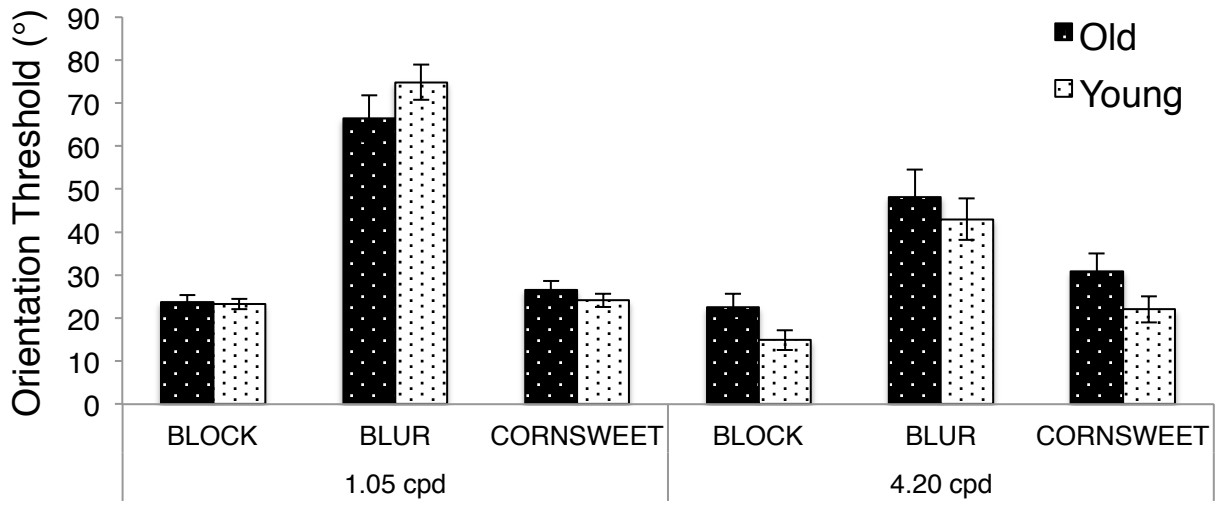


Figure 6.8. Average orientation contrast thresholds for the Segmentation of texture stimuli. Error bars represent the standard error of the mean.

Table 6.1
Summary of ANOVA results on orientation contrast thresholds.

	df	*F*	*P*
Age	1,20	0.852	0.367
LSF	1,20	22.022	**<0.001**
Profile	2,40	151.602	**<0.001**
Age × LSF	1,20	5.680	**0.027**
Age × Profile	2,40	1.542	0.226
LSF × Profile	2,40	14.614	**<0.001**
Age × LSF × Profile	2,40	0.387	0.682

LSF=luminance spatial frequency

The main effect of luminance spatial frequency shows that on average, orientation thresholds for the low (1.05cpd) luminance spatial frequency were higher than the high (4.20cpd) luminance spatial frequency (High LSF=30.3°; Low LSF=39.6°). The main effect of profile shows that the Block profile had the lowest orientation thresholds (21.1°) followed by the Cornsweet profile (25.9°), and the Blur profile had the highest orientation thresholds (57.9°) [Bonferroni corrected post hoc comparisons showed that for Block vs. Cornsweet, p=0.001, while p<0.001 for all other comparisons]. Additionally, a significant luminance spatial frequency and profile interaction was found, where Bonferroni corrected pairwise comparisons (see Figure 6.9 left) showed that for the Block and Cornsweet profile, the luminance spatial frequency does not affect orientation thresholds. However, for the Blur profile, orientation thresholds are lower for the high luminance spatial frequency compared to the low luminance spatial frequency. A luminance spatial frequency and age interaction was also found, where Bonferroni corrected pairwise comparisons (Figure 6.9 right) showed that younger participants have lower orientation thresholds for the high luminance spatial frequency compared to the low luminance spatial frequency, while older participants shows no difference between the luminance spatial frequency.

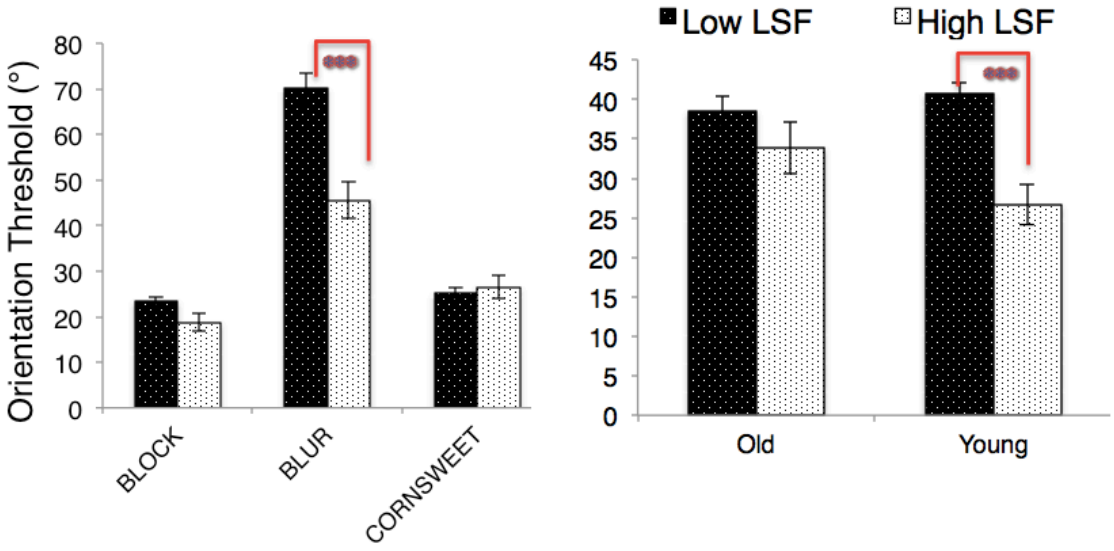


Figure 6.9. Left graph shows the luminance spatial frequency and profile interaction for the Segmentation Task. Right graph shows the luminance spatial frequency and age interaction for the Segmentation Task.

6.2.3 Modeling

The modeling of the Integration Region (IR) was done in the same manner as described in Section 3.5. However, when modeling the results for this experiment, the orientation thresholds for the three profiles could only be re-expressed in terms of the different mean orientations for raster units up to N=4, instead of N=6, owing to the smaller 8×10 figure grid.

The IRs are calculated for each group of participants at both luminance spatial frequencies (Figure 6.10 shows the average IR size). A 2 (luminance spatial frequency) × 2 (age group) × 3 (orientation profile) Mixed-Design ANOVA only found a significant main effect of luminance spatial frequency $F(1,19)=11.364$, $p=0.003$, showing that the IR size is larger with high luminance spatial frequency ($M_{LSF}=1.911$; $M_{HSF}=1.451$). No main effect of age $F(1,19)=3.204$, $p=0.098$ or an interaction between age and luminance spatial frequency $F(1,19)=0.038$, $p=0.847$ was found.

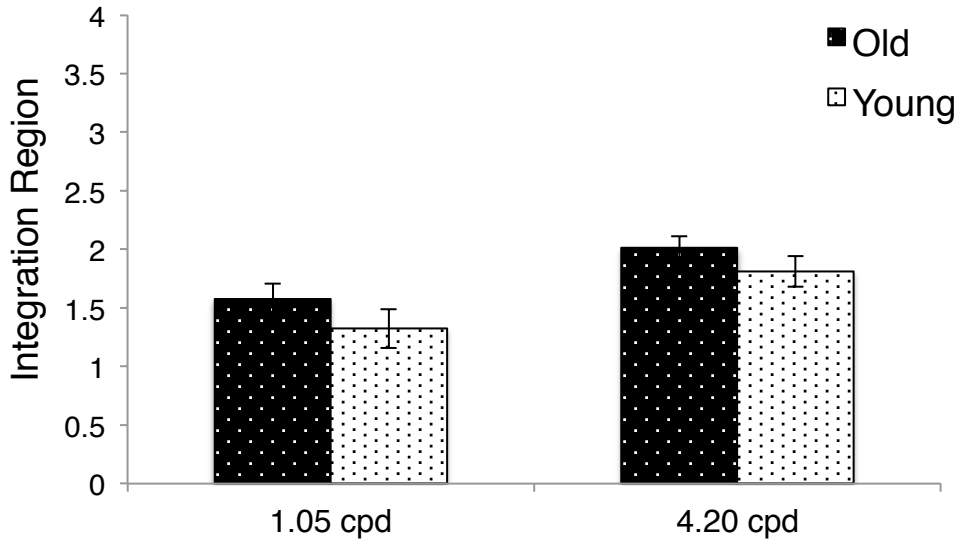


Figure 6.10. Average Integration Region for the older and younger participants at both luminance spatial frequencies. Error bars represent the standard error of the mean.

6.3 Experiment 6: Scale Invariance

The results of Experiment 5 show an interesting trend in the results. Performance (in terms of thresholds) for the Orientation Discrimination task and Segmentation task show an expected outcome whereby older adults tend to have poorer performance for high luminance spatial frequencies (4.20 cpd). Despite this, the IR from the modeling shows no difference in integration size between old and young participants, nor an interaction between age and luminance spatial frequency. We did however find a difference in IR size for the different luminance spatial frequencies, but it was in the opposite direction of what we would predict based on Kingdom and Keeble (1999) (see Section 6.4 for discussion).

This current study will employ the same task and stimuli as Experiment 5, but it will be conducted at different viewing distances. The purpose of doing this experiment is 2-fold: 1) we want to investigate why the IR size is contradictory to Kingdom and Keeble (1999), and 2) we want to investigate why the IR is not affected by age. To elaborate, Kingdom & Keeble

(1999) found that the orientation modulation function was scale invariant – that is when viewing distance is changed, performance stayed the same when expressed in terms of physical units on the screen. Therefore one would expect that the mechanism that does the task is scale invariant. By doing this study, we will be able to test if the mechanism that is doing the task in our study is also scale invariant. It is possible that the difference in results is because a completely different mechanism is being employed.

Regarding the IR and age, we know that older participants are having difficulties with the high luminance spatial frequency, but this does not seem to impair their IR size. This could mean that as long as performance is above a certain level (i.e. they can do the task), IR is not affected. Therefore, in this study, when the viewing distance is increased, it will be harder to see the stimuli and perform the task (i.e. the participants' performance will be impaired just like the older participants), but the IR size should stay the same.

6.3.1 Methods

6.3.1.1 Participants

Four observers (DK, SKS, PE and AI) participated in the experiment. DK (age = 55) and SKS (age = 26) were experienced psychophysical observers, while PE (age = 22) and AI (age = 22) were naïve observers. All participants had corrected-to-normal vision.

6.3.1.2 Display

The display was the same as described in Section 3.2, with the exception that the pixel resolution was 1.8 arcminutes per pixel when viewed from a distance of 72cm.

6.3.1.3 Stimuli

The stimuli used for the Segmentation Task and Orientation Discrimination Task was exactly the same as Experiment 5 (Section 6.2.1.3).

6.3.1.4 Procedure

The experimental procedure was once again exactly the same as Experiment 5 (6.2.1.4). However, in Experiment 5, the participants performed the task at only one viewing distance (54cm, which is equivalent to 72cm in this computer setup). In the present study, the participants performed the task at a total of three viewing distance: 72cm (which has been made identical to Experiment 5), 36cm (1/2 the distance), and 288cm (×4 the distance). Participant SKS also performed the task at two additional viewing distances: 18cm (1/4 the distance), and 144cm (×2 the distance). For all participants, the Orientation Discrimination Task was performed first, followed by the Segmentation Task. For each of the tasks, blocks of different viewing distances were randomised.

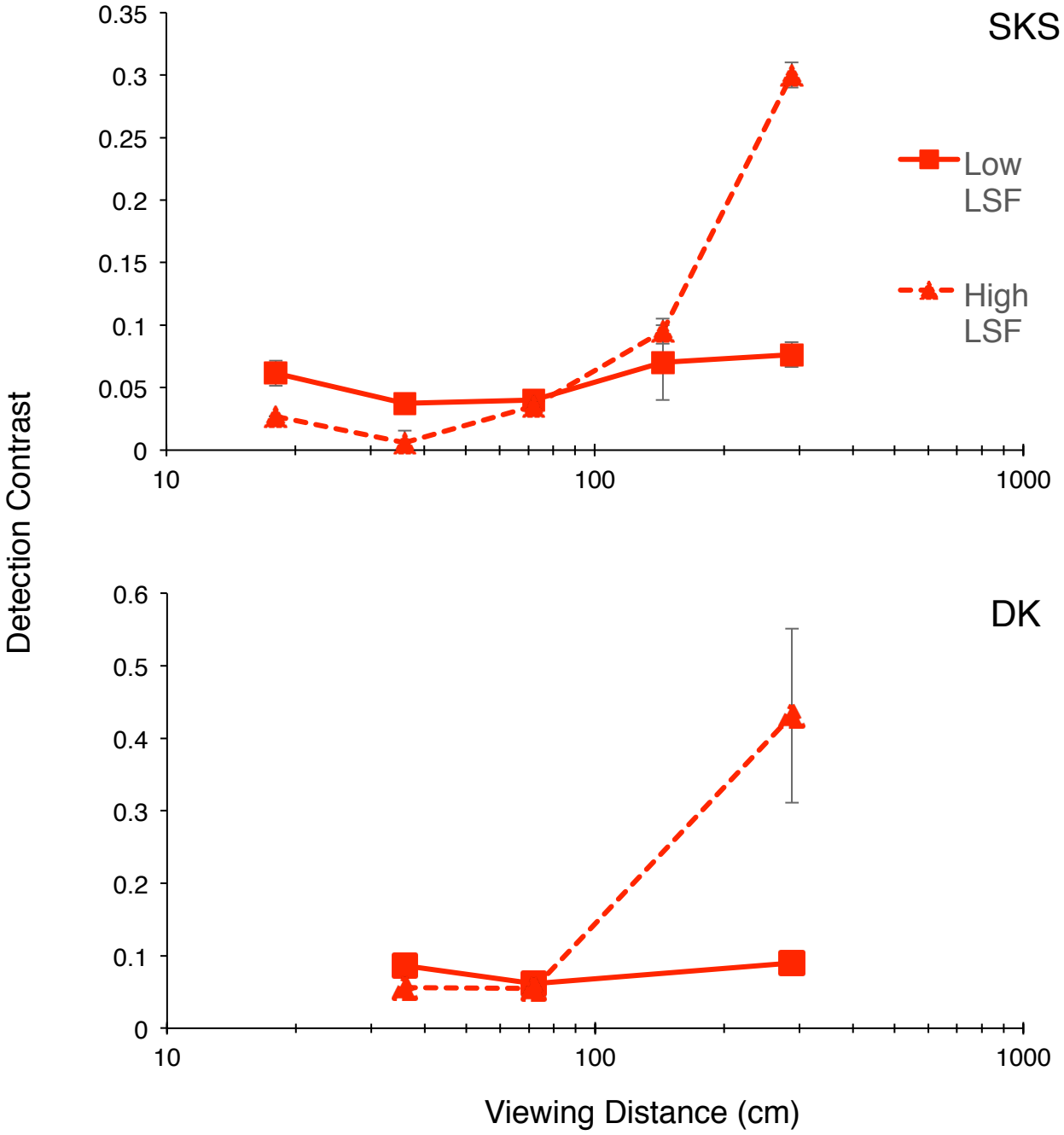
6.3.2 Results

In this study, at the original viewing distance (72cm, which is equivalent to Experiment 5), the high luminance spatial frequency was at 4.20cpd, and the low luminance spatial frequency was at 1.05cpd. However, as the viewing distance changes, the luminance spatial frequency in terms of cpd also changes. As such, from now on, high and low LSF will refer to respectively 1.05cpd and 4.20cpd at the original viewing distance.

6.3.2.1 Orientation Discrimination of a Single Gabor Element

Contrast thresholds for the four participants are plotted in Figure 6.11 below, with Figure 6.12 showing the mean data across participants. For participants SKS and DK, contrast

thresholds remain relatively constant with an increase in viewing distance for the Low LSF (changes in contrast threshold was never more than 1.4× higher). For participants PE and AI, contrast thresholds at the Low LSF increases by 6.9× only for the viewing distance of 288cm. For the High LSF, all participants show very little change at the closer viewing distance, with an average increase of 1.4× from the 36cm viewing distance to the 72cm viewing distance. However, there was an increase in contrast threshold of 18.8× for the 288cm viewing distance compared to the 72cm viewing distance. Participant AI was unable to produce a physically meaning contrast threshold (i.e. below 1) for the 288cm viewing distance of the High LSF.


SKS
DK
Low LSF
High LSF
Detection Contrast
Viewing Distance (cm)
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
0.6
0.5
0.4
0.3
0.2
0.1
0
10
100
1000

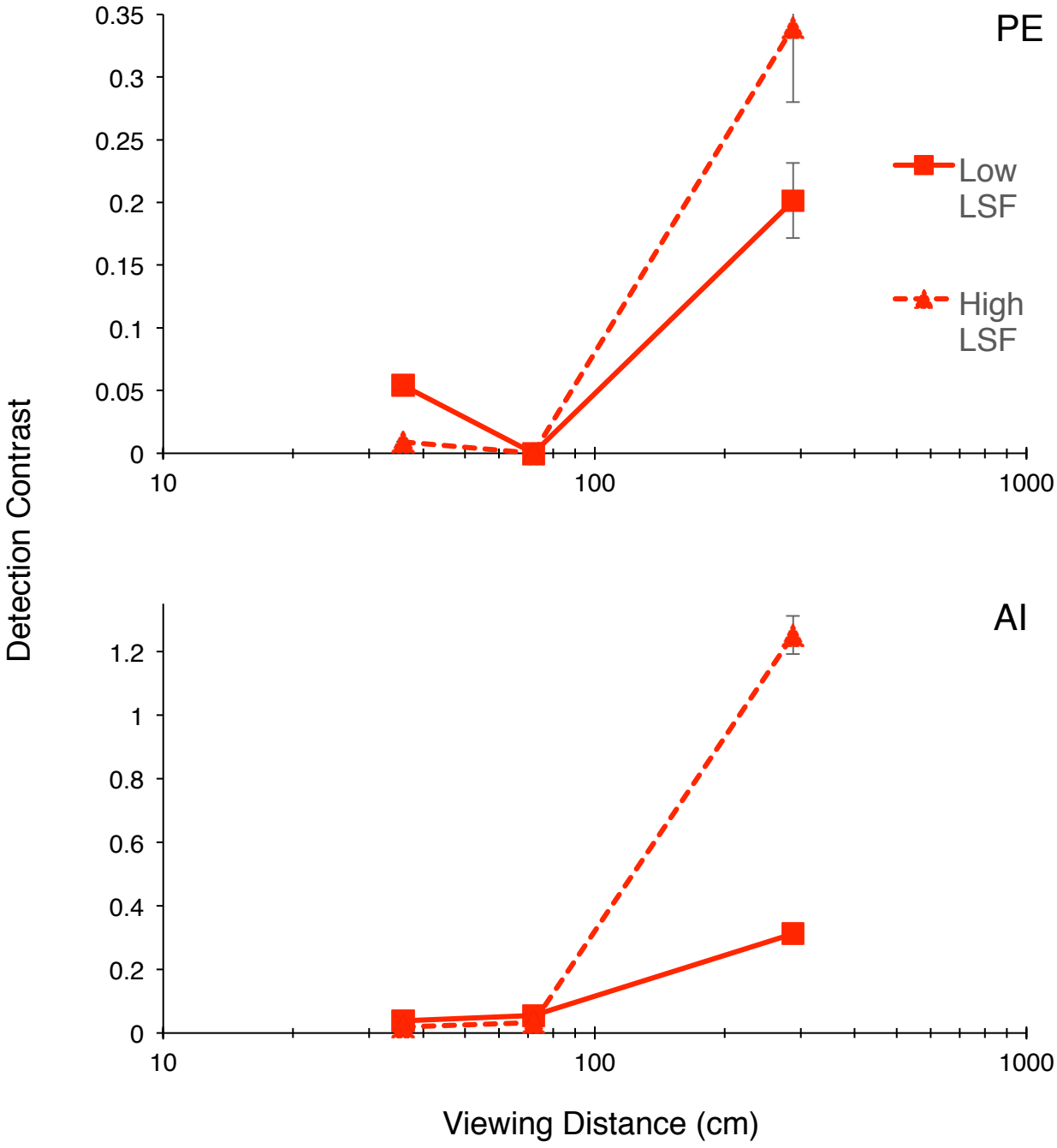


Figure 6.11. Contrast thresholds for the Orientation Discrimination of a single Gabor element is plotted as a function of viewing distance is. Error bars represent the 67% confidence interval, and where no error bars are seen, the values are small and the data point obscures the bars.

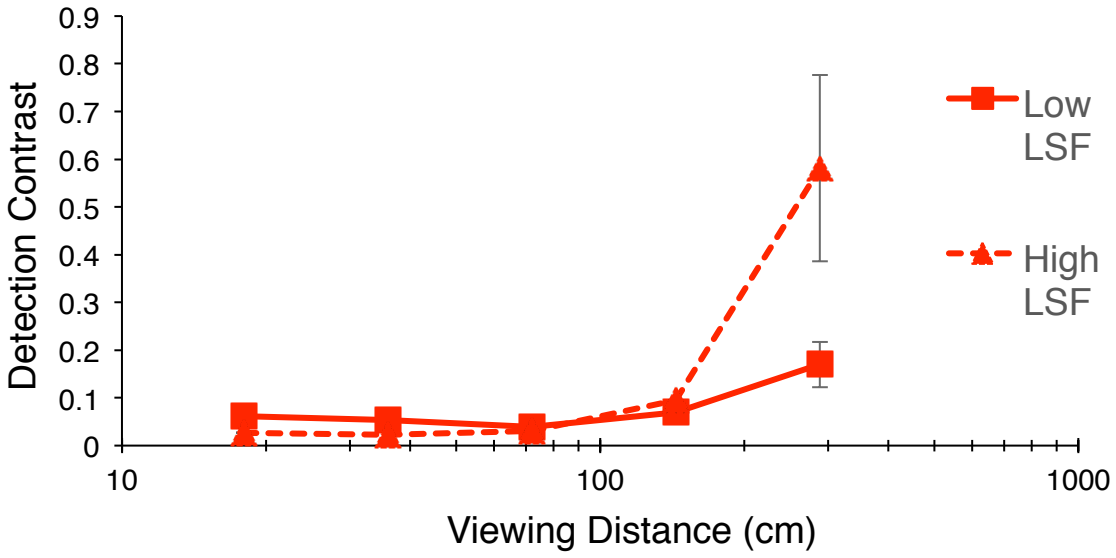
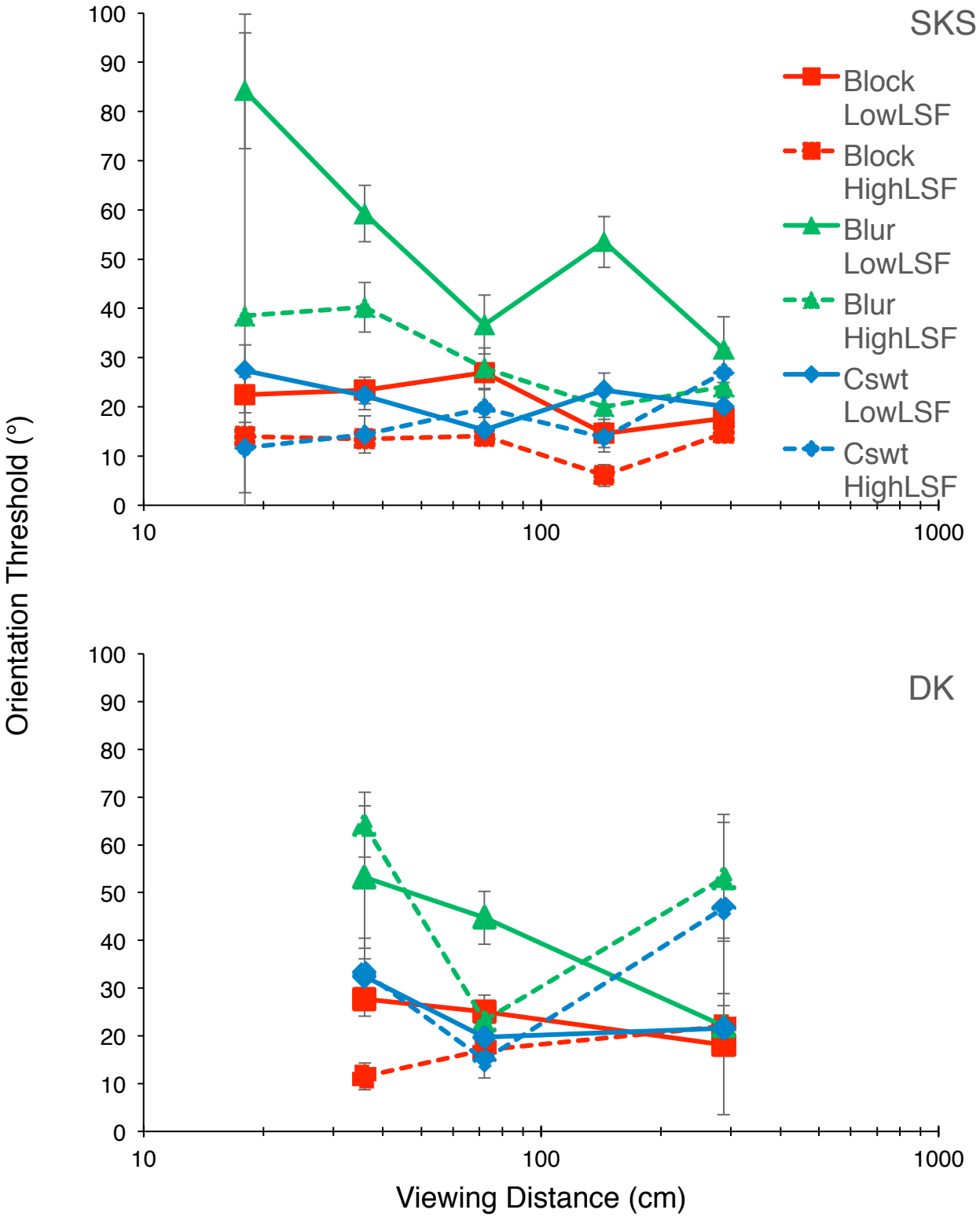


Figure 6.12. Average contrast thresholds of all participants (*n*=4) for the Orientation Discrimination of a single Gabor element is plotted as a function of viewing distance is. Error bars represent the standard error of the mean, and where no error bars are seen, the values are small and the data point obscures the bars.

6.3.2.1 Segmentation of Texture Stimuli

Weibull functions were fitted to the psychometric functions, and thresholds at 70% correct level was used. The 70% correct level was used for two reasons – firstly, for consistency with Experiment 5, and secondly, the naïve observers also had difficulty producing physically meaningful thresholds for all the various conditions at 82% correct. The orientation thresholds for the various conditions for all participants are plotted in Figure 6.13, with the average response across participants being plot in Figure 6.14. Inspection of Figures 6.13 and 6.14 do not show very conclusive trends in the data. As such, the modelling of the results (see Figure 6.15 in Section 6.3.3) would be best used to interpret the findings of this experiment.


SKS
DK
Block LowLSF
Block HighLSF
Blur LowLSF
Blur HighLSF
Cswt LowLSF
Cswt HighLSF
Orientation Threshold (°)
Viewing Distance (cm)
0
10
20
30
40
50
60
70
80
90
100
10
100
1000

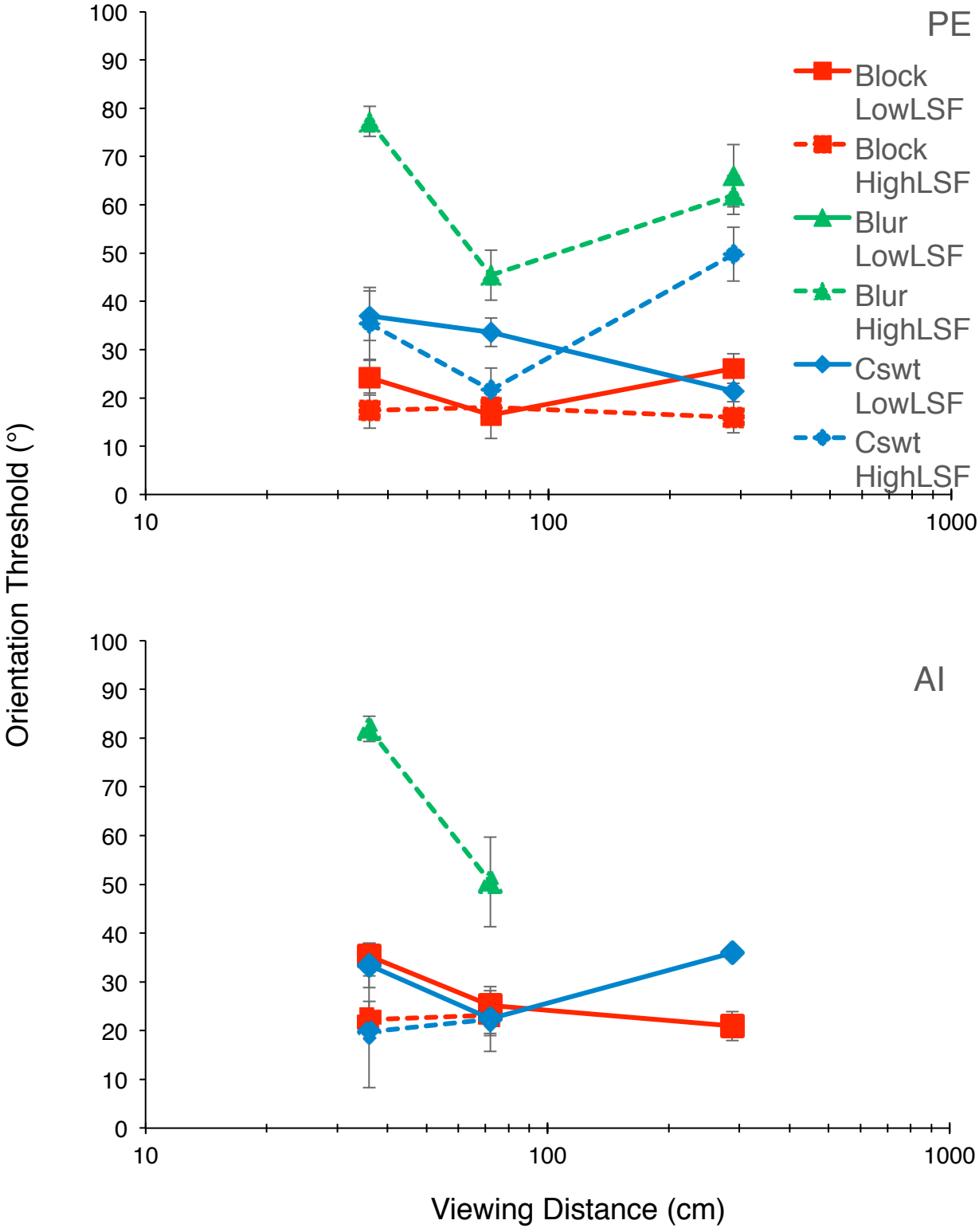


Figure 6.13. Orientation contrast thresholds are plotted for the Segmentation of the texture stimuli as a function of viewing distance. Error bars represent the 67% confidence interval. Dashed lines for High LSF, and solid lines for Low LSF. Where physically meaningful thresholds were not achievable, the data was not plotted.

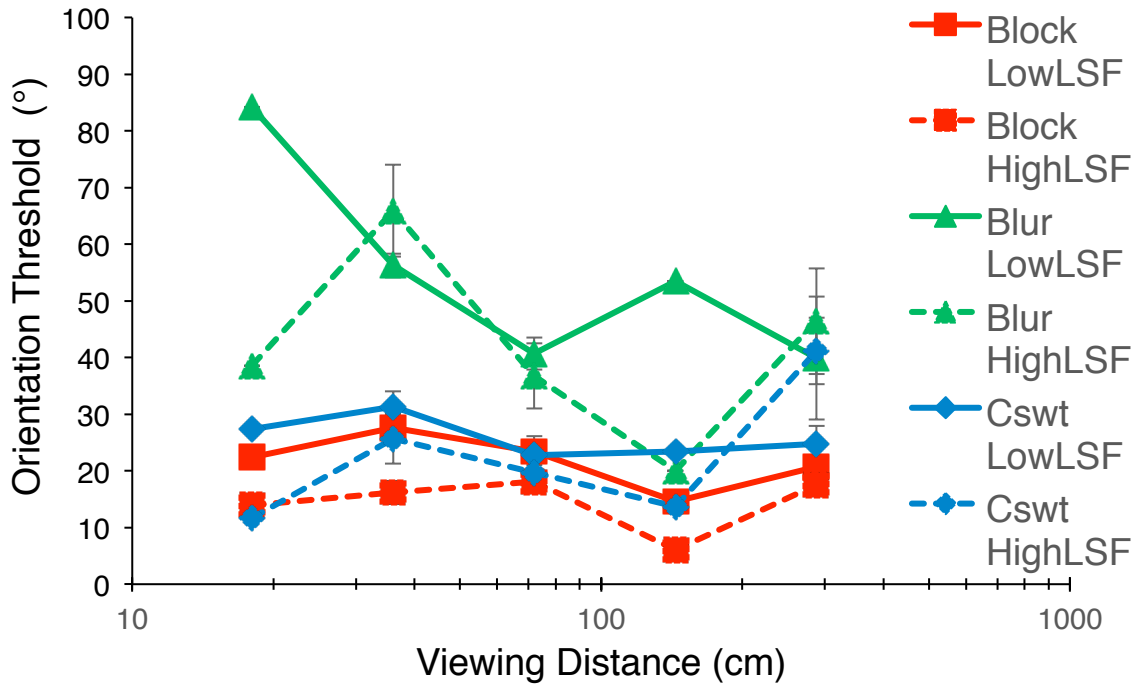
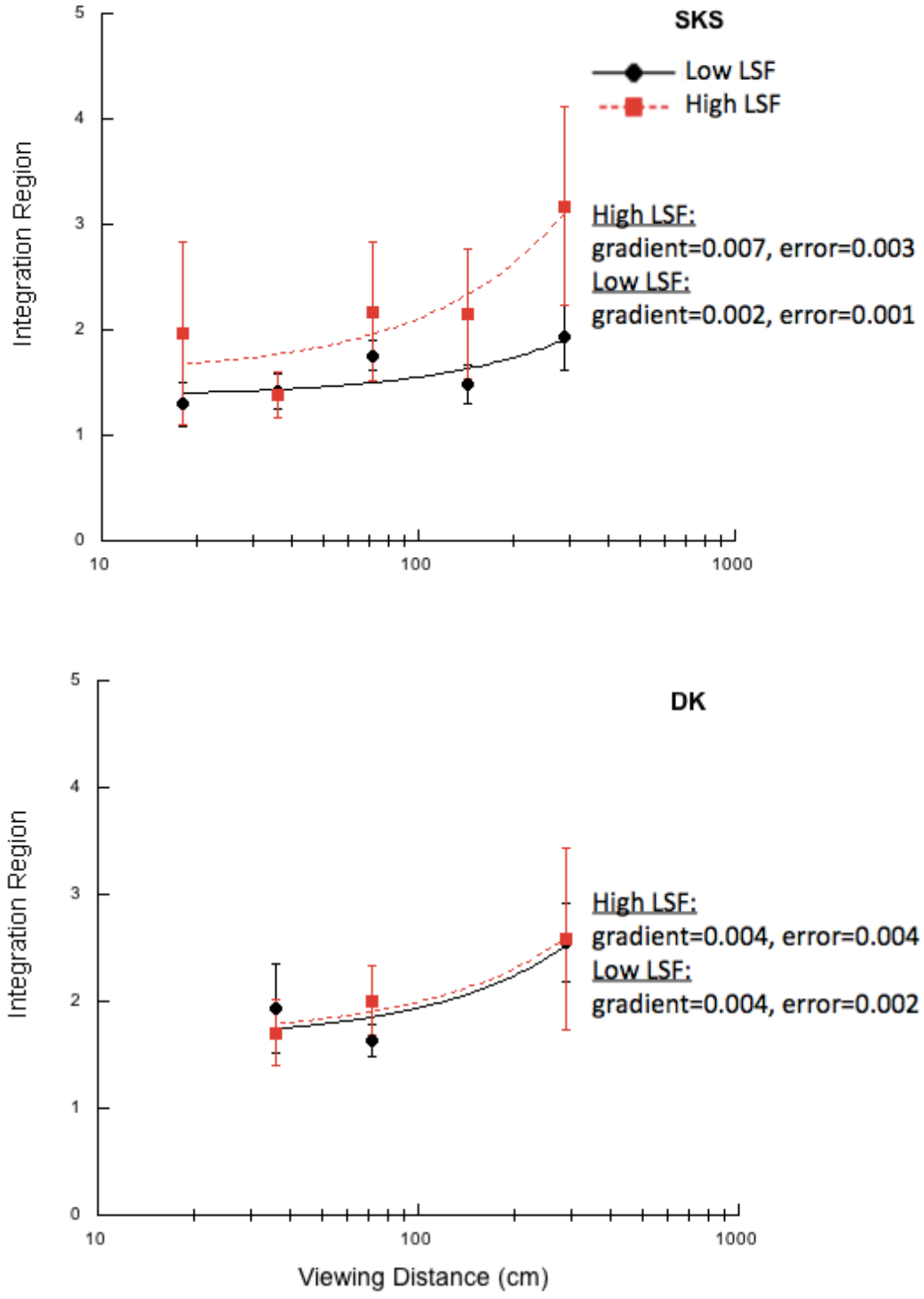


Figure 6.14. Average orientation contrast thresholds of all participants (*n*=4) are plotted for the Segmentation of the texture stimuli as a function of viewing distance. Error bars represent the standard error of the mean. Dashed lines for High LSF, and solid lines for Low LSF. Where physically meaningful thresholds were not achievable, the data was not plotted.

6.3.3 Modeling

The modeling of the Integration Region (IR) was done in the same manner as described in Section 3.5. Figure 6.15 shows the IR of the four participants as a function of viewing distance for both High and Low luminance spatial frequencies (LSF). In accordance with Experiment 5, the High LSF appears to have a slightly larger IR compared to the Low LSF. Furthermore, viewing distance has very little affect on the IR as physical units on the screen. It is possible to argue that there is a slight increase in IR, especially for the High LSF. However, a key point to note is that the viewing distance changes by a factor of 8 (and 16 for participants SKS), but the IR changes at most by a factor of 1.5 (or 1.6 for participant SKS). Besides that, we fit a linear line of best fit to the data to determine the gradients of the slope. Disruption of scale invariance should produce slopes with gradients greater than 0. In

Kingdom and Keeble (1999), when the authors found evidence of partial disruption to scale invariance, the gradients of the slopes were between –0.46 and –0.76, whereas they were between –0.06 and –0.13 when the results were scale invariant (Figures 6, 8, and 9 from Kingdom & Keeble, 1999). Our results (see Figure 6.15) show that the slopes are nearly flat, with gradients between 0.001 and 0.008. Therefore, it is reasonable to claim that the results here are approximately scale invariant i.e. the perceptual properties of the textures do not change with viewing distance.


SKS
Low LSF
High LSF
High LSF:
gradient=0.007, error=0.003
Low LSF:
gradient=0.002, error=0.001
Integration Region
0
1
2
3
4
5
10
100
1000
DK
High LSF:
gradient=0.004, error=0.004
Low LSF:
gradient=0.004, error=0.002
Integration Region
Viewing Distance (cm)

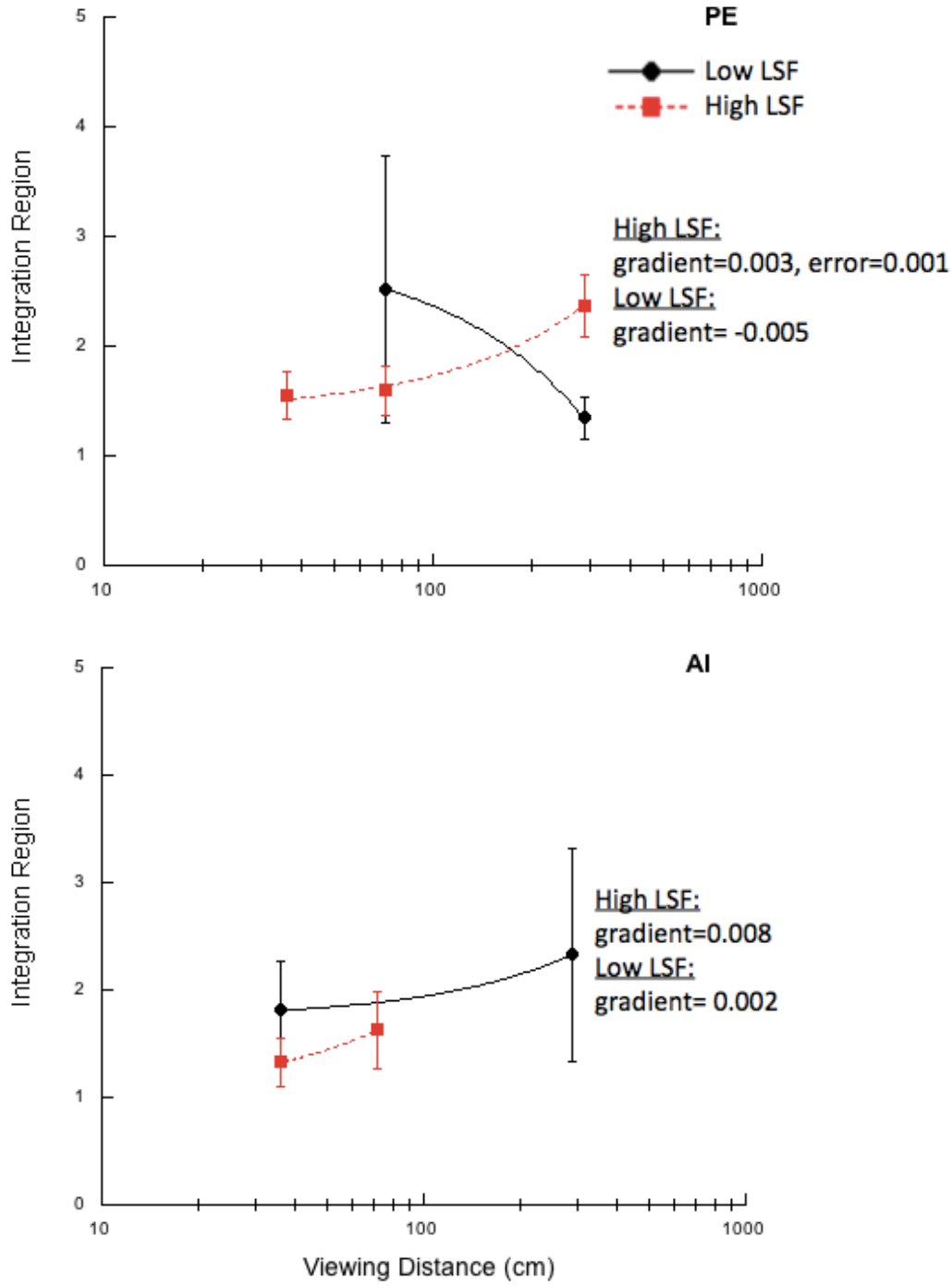


Figure 6.15. Integration Region of the four participants plotted as a function of viewing distance. The error bars represent the standard deviation, which was obtained by bootstrapping the data. A line of best fit was added to determine the gradient of the slope. Where shown, the error represents the error estimate of the gradient. This was determined by weighting the data from the error associated with the integration size (error bars on data points).

6.4 Discussion

In the pilot study, we demonstrated how the spatial frequency content in a grating affects orientation thresholds (degree of orientation contrast needed to perform the task) in a texture segregation task. This was also studied by Landy and Bergen (1991), but they only tested a few spatial frequencies. In our study, we tested a greater range of spatial frequencies, and our results are in partial accordance with Landy and Bergen (1991). Partial because in their paper, they claim that performance improves with higher spatial frequencies (which we also found), but since we tested a greater range of spatial frequencies, we also know that performance gets worse at spatial frequencies greater than 5.7cpd. Additionally, we controlled for the confounding effects of visibility by measuring participants' contrast threshold first, and subsequently (in the Matched Contrast study) increasing the contrast of the stimuli by fixed decibels thereby equating the visibility of the stimuli. Orientation contrast thresholds were very similar for both the Matched and Maximum Contrast study, suggesting that texture segregation performance was not dependant on the visibility of the Gabor elements. Thresholds for both studies (Maximum and Matched Contrast) were very similar, suggesting that a second-order mechanism is receiving inputs of different intensities from the first-order luminance spatial frequency, which appears to be independent of contrast.

In Experiment 5, we investigated how old and young adults performed in an Orientation Discrimination task and Segmentation task. The first thing we shall discuss is the contrast threshold for the Orientation Discrimination task. Owsley et al., (1983) found that for 1cpd gratings, sensitivity for old and young age groups were the same, and only at higher spatial frequencies (2cpd and above) did sensitivities decrease for older adults compared to younger adults. In contrast, our results show that older adults consistently had lower sensitivity

(higher thresholds) compared to young adults. However, our results do show that older adults have decreased sensitivity when detecting Gabor elements of higher spatial frequency compared to low spatial frequency. Young adults show no difference in sensitivity to high and low spatial frequencies. This difference in results could be because Owsley's study differed from ours in a few crucial ways: we used Gabor elements while they used sine wave gratings, and their stimuli were presented at fovea, while ours could appear anywhere on a circle centred on the fixation. Our contrast thresholds were more in line with Delahunt et al., (2008), who found that older adults had higher contrast thresholds than younger adults, regardless of spatial frequency. Nevertheless, they found that contrast thresholds were higher for the 4cpd Gabor element than the 1cpd Gabor element (we found no difference), and they did not find the interactive effect of age and spatial frequency like us.

One prominent result from the Segmentation task was that there was no overall difference in thresholds between young and old adults i.e. old and young adults perform equally as well. Some studies (Habak & Faubert, 2000; Faubert 2002; Tang & Zhou, 2009; Bertone et al., 2011) claim that older adults have significantly impaired performance when processing second-order stimuli. However, our study used stimuli that would require second-order processing, and yet older adults thresholds were similar to the younger adults. This difference in results could possibly be due to the different type of stimuli used in the aforementioned studies and our study. Those authors used contrast-defined second-order stimuli of sine wave gratings (Habak & Faubert, 2000; Tang & Zhou, 2009) and Landolt-Cs (Bertone et al., 2011), while our study used orientation texture-defined second-order stimuli. This indicates that older adults do not necessarily have impairments processing all second-order stimuli, perhaps just contrast-defined stimuli.

In fact, studies have shown that for motion stimuli (also requiring second-order processing), older adults do not always perform worse than younger adults. Depending on the stimulus parameters, older adults perform equally as well (Snowden & Kavanagh, 2006) or sometimes even better (Betts et al., 2005; Hutchinson et al., 2014) than younger adults. These findings as well as the results presented in this chapter argue against Faubert's theory that second-order processing faces greater deficits in the elderly. As far as complexity of the task goes, one could argue that the elderly participants in our study were able to perform well as the stimulus duration (1000ms) was suitably long enough to decrease the difficulty of the task. However, Bertone et al., presented the stimulus for 2000ms, and the elderly participants still performed markedly worse than young participants for the second-order contrast-defined stimuli.

Besides that, Snowden and Kavanagh (2006) also found that older participants were worse than younger participants at judging the direction of motion of a sine wave grating for all spatial frequencies tested (0.5 – 4cpd). This is in contrast to the results of the Segmentation task, wherein an interactive effect of age and spatial frequency was found. This interaction is explained by young adults having lower thresholds for the 4cpd compared to the 1cpd spatial frequency, while older adults show no difference in thresholds for both spatial frequencies. As a matter of fact, post hoc comparisons between young and old participants showed no difference in thresholds between the low and high luminance spatial frequencies.

Both Betts et al., (2005) and Hutchinson et al., (2014) found patterns of results whereby older participants performance remained constant with changing stimulus size, while younger participants' performance declined with motion stimuli. For the textured-stimuli used in our study, increasing spatial frequency did not affect segmentation performance in the elderly, but it did for the young adults, where performance improved for high spatial frequency

stimuli. Betts et al., (2005) suggested that the reason older people perform better than younger people with larger stimuli is due to reduced centre-surround interaction. Though according to Lamme (1995), the interactions between the surround field and the classical receptive field are crucial for figure-ground segregation. These two studies would suggest that older people would exhibit deficits in any type of orientation-based segregation task, but we do not see that. In fact, in most cases, older participants have very similar thresholds to the younger participants. Therefore, our results would suggest that centre-surround interaction is not necessarily reduced in older participants.

Other studies (Hadad, 2012; Casco et al., 2011) have also proposed that older people have reduced inhibitory processes – which causes increased internal noise according to Hadad (2012), and affects the suppressive mechanism according to Casco et al (2011). If older people do in fact have reduced inhibitory processes, we would actually expect that in our study, older people would have performed worse than younger people as increased internal noise and an inability to supress background information would have increased orientation thresholds when performing the task. For instance, studies (e.g. Rossi, Desimone, & Ungerleider, 2001; Lamme, Rodriguez-Rodriguez, & Spekreijse, 1999) have found that contextual modulation (where there is increased neural response to a figure compared to the background) plays a role in texture segregation. Therefore, if older adults had diminished abilities to supress background information, performance should have been impaired.

In contrast to studies that suggest reduced inhibition in the elderly, Wang et al., (2018) found evidence of stronger lateral inhibition in older adults. Lateral inhibition is when an excitatory neuron inhibits the activity of its neighbouring neurons. In figure-ground segmentation, where edge extraction is important, lateral inhibition plays a very important role (e.g. Roelfsema et al., 2002). This is because neurons at the edge receive less lateral inhibition

than the centre. Therefore, if older adults do have stronger lateral inhibition, their performance should have been better than the younger adults.

The results of our study is more in line with those that suggest that older people do not have impairments in encoding orientation information (Delahunt et al., 2008; Govenlock et al., 2009; Casco et al., 2017; Wang et al., 2018). This could explain why older adults performance is not that different than younger adults in our study. In fact, Delahunt et al., found that when stimulus contrast was expressed in terms of multiple of contrast threshold, performance for older and younger participants did not differ. This might explain the interactive effect we found between age and spatial frequency for the Segmentation task. Older participants' contrast threshold for the 4.20cpd luminance spatial frequency (LSF) was higher than the 1.05cpd LSF, while younger participants showed no difference in contrast threshold for both LSFs.

Our results, while in line with Casco et al., (2017)'s preservation of orientation encoding in adulthood, does not fit in with the inefficient decoding (extracting sensory information from the brain to make a decision) in adulthood idea. According to them, neurons respond to both the target and the distractors, and reduced inhibition makes the correct choice harder to be made in a 2AFC task, thus resulting in higher rates of false alarms. However, our study employed a 2AFC, where the figure had to segregated from the background (which serves as a distractor), and the participant had to make a decision as to the shape of the figure. We found that older participants were able to perform this task – and it most cases, perform equally as well as the young adults, suggesting that older adults do not have impaired decoding.

Interestingly, despite the fact that younger participants show an advantage over the older participants in processing high spatial frequency content, when we modelled the data in terms

of an Integration Region (IR), we see no effect of age, nor an interaction of age and spatial frequency. We might expect that older adults would have a larger IR (Betts et al., 2005), or even a smaller IR (Hutchinson et al., 2014), but it is rather unexpected to see no difference here. There are two possible reasons that we do not see a difference in IR between old and young adults: 1) the study by Hutchinson et al., (2014) and Betts et al., (2005) used dynamic motion-defined stimuli, which could possibly not reflect any of the same mechanisms for our static stimuli, and 2) the fact that older people perform better than younger people in those studies depend on the actual stimulus size being large or small, while the IR is the size of a possible neural mechanism i.e. the size of the second-order filter that we estimated from computational model. As such, the similar IR sizes for old and young adults participants found in our study are more in line with that of Hadad (2012) and Casco et al., (2011) who found that integrative mechanisms for contour integration are preserved in old age. This required further investigation on our part (Experiment 6).

Since we found an interactive effect between age and spatial frequency for the Segmentation task but not the IR, it could mean that the IR operates as an all-or-nothing mechanism, wherein no matter how difficult the task (like the older participants faced), the IR will not change. The modeling of Experiment 6 gives support to this theory. When we had participants perform the task at different viewing distances, the task was made much harder with further viewing distances. Despite this, the IR barely changed with viewing distance. In fact, based on the contrast sensitivity of participant AI (High LSF at 288cm), AI was not able to see the grating at that spatial frequency and viewing distance. Therefore, when performing the Segmentation task, AI could not produce physically meaningful thresholds or an IR. This could explain what is happening with the older adults too. As long as the gratings were visible to them, they could perform the task (even if they were slightly poorer than their younger counterparts), and IR was not affected. If we return to the idea of an all-or-nothing

mechanism that integrates information for visual processing, for all the participants (older and younger) the IR works in the same manner.

To discuss the non age-related changes in the data, a few aspects are most prominent. First, like Landy and Bergen (1991) and the pilot study conducted, we found that performance was better for the high spatial frequency. Second, an effect of profile was also found (Block > Cornsweet > Blur), which is in line with our previous findings (see Chapters 4 & 5). Third, threshold performance shows that while spatial frequency content has limited effect on the Block and Cornsweet profile, there is an advantageous effect on the Blur profile for the high spatial frequency. Finally, the modeling of the data showed that the IR is larger for the high spatial frequency compared to the low spatial frequency.

Our IR results for the spatial frequency are in contrast to what we might expect from Kingdom and Keeble (1999). The authors posit a two-stage model, wherein the first-stage filters are tuned to selectively receive input about the luminance spatial frequency, which then feeds into the second-stage texture filters. They suggest that there is a range of sizes for the second-stage filters, which are selective based on the luminance spatial frequency of the first-stage input. In sum, small second-stage filters will be used for high luminance spatial frequency from the first-stage, and large second-stage filters for low luminance spatial frequency input.

The difference in results seen between our study and that of Kingdom and Keeble (1999) could be because it is not necessarily the same mechanism that is being used. In Experiment 6, we tested whether the mechanism used in our study is also scale invariant, as shown by Kingdom and Keeble (1999), or was not, implying that a different mechanism was being employed. Our results show that the IR is scale invariant, which makes the results more puzzling, but it does not necessarily mean that the mechanism used is the same. There are

quite a few dissimilarities between our study and the one done by Kingdom and Keeble (1999), which could account for the difference in results seen. Primarily, their study measured orientation modulation functions (OMF), which was the orientation modulation needed to perform the task (2-interval forced choice task) at different texture spatial frequencies. Subsequently, for each viewing distance tested, they used the OMF data to find the texture spatial frequency that had the lowest threshold (TSF_{min}). When the TSF_{min} was plotted as cycles per screen instead of cycles per degree for the different viewing distance, the slope was nearly flat, demonstrating scale invariance. While the TSF_{min} (cycles per screen) and the IR in our study were expressed in terms of physical units on the screen, the IR was calculated based on the integration size in which the thresholds of the three different profiles used are similar.

Finally, a manner in which we can interpret the results is like that of Del Viva and Agostini (2007), who put forth the idea that for contour integration there are two separate mechanisms for large and small contour spacing, and these mechanisms are affected differently by the aging process. In the case of orientation-based texture segmentation, there could be mechanisms of different sizes for the different spatial frequencies (as we found that IR for high luminance spatial frequency is larger), and for the spatial frequencies we tested here, the integrative mechanisms show no change with age.

In sum, the results presented in this chapter show no evidence that older adults have impairments in processing second-order stimuli as suggested by Habak and Faubert, 2000; Faubert, 2002; Tang and Zhou, 2009; Bertone et al., 2011. Nevertheless, unlike studies that have observed improved performance in the elderly compared to young adults (e.g. Allen et al., 2010; Hutchinson et al., 2014; Betts et al., 2005), our results show that young and old

adults perform equally as well in a texture segmentation task. Thus, the mechanism that mediates texture segmentation appears to be preserved in old age.

Chapter 7

Introduction to Eye Movement Research

7.1 Introduction To The Eye

Considering how half the work presented in this thesis uses eye tracking measures, this section is going to very briefly cover the physiology of the eye. This is mostly to lay the foundation of what we will cover in the next few sections, which are the Characteristics of Eye Movements (Section 7.2) and Measuring Eye Movements (Section 7.3).

Visual information from the environment is received by the eyes, and carried through the visual pathways for processing in the visual cortex and other parts of the brain. The eye is a fluid-filled sphere that is enclosed by three layers of tissue – the sclera (outer-most layer), uveal tract (middle layer), and retina (innermost layer). Within the retina, neurons called photoreceptor cells are capable of converting light to visual signals due to their sensitivity to light. Images are formed on the retina by the cornea and lens, which are able to refract (bend) light so that focused images are formed on the photoreceptors of the retina. The axons of the retinal ganglion cells, which form the optic nerve, carry the information from the retina to the rest of the brain (Purves et al., 2012).

The formation of inverted images on the retina is achieved when light passes through the pupil (small hole in the iris). The light is brought to focus by the cornea and lens. The cornea is the transparent surface in the front of the lens, which is accountable for the majority of the refraction needed. The lens is another transparent structure in the eye that is able to refract light, and while it does this at a lesser extent than the cornea, the refraction produced by the lens is adaptable (Rodieck, 1998). This ability to change the refractive power of the lens is

known as accommodation, and this response is achieved via the ciliary muscle, which changes the shape of the lens. When viewing objects far away, the lens takes a thin and flat shape by relaxation of the ciliary muscle to acquire low refractive power. When viewing objects nearby, the lens becomes thicker and rounder by contraction of the ciliary muscle to amass greater refractive power (Purves et al., 2012).

Within the retina, there are two types of photoreceptor cells – rods and cones – that are specialized cells to perform different functions. The cone cells perform optimally at bright light (low sensitivity to light), and have high acuity (high spatial resolution). The three types of cone cells are what allow humans to perceive colour. The rod cells on the other hand are able to function at low light conditions (high sensitivity to light), but have low acuity (low spatial resolution) (Rodieck, 1998). Rod cells do not play a role in colour vision, but are entirely responsible for night vision in humans. Thus, at extremely low light levels, only rod cells are activated (scotopic vision). As light levels increase, cone cells start contributing to perception (mesopic vision), and at high light levels such as indoor lighting or sunlight, cone cells are entirely responsible for visual perception as the contribution from rod cells is saturated (photopic vision) (Purves et al., 2012).

Other than sensitivity to light and spatial resolution, the distribution of the rod and cone cells also varies across the surface of the retina. The number of rod cells (approximately 90 million) in the retina far outweighs the number of cone cells (approximately 4.5 million) despite the fact that perception in bright light conditions, which is the condition we are most accustomed to, is mediated entirely by the cone system. However, in the fovea, which is a small depression/pit in the retina that measures approximately 1.2 millimetres in diameter, the number of cone cells far exceeds the number of rod cells. In fact, at the central 0.3 millimetres of the fovea, there are no rod cells present (Purves et al., 2012).

This densely packed area with cone cells allows extremely high levels of visual acuity at the fovea. The further away the image of an object is from the fovea, the lower the acuity of the object. A displacement of 6° from the fovea decreases acuity by 75% (Palmer, 1999). As such, eye movements are constantly being made to bring regions of interest to the centre of the fovea where acuity is at its highest. These eye movements are made using three antagonistic pair of muscles – superior and inferior rectus muscle (vertical movement), superior and inferior oblique muscle (torsional movement), and lateral and medial rectus muscle (horizontal movement) (Purves et al., 2012). Details regarding the types of eye movements made are discussed in detail below.

7.2 Characteristics of Eye Movements

In this section, we will first give a brief overview of all the different types of eye movements, even though we are only interested in saccadic movements in our experiments. Subsequently, we will go into more details about saccadic eye movements, in particular saccadic amplitudes and latencies, as these are measures we are primarily interested in. Finally, in the interest of understanding **why** and **how** we make eye movements, we will discuss the control of eye movements.

7.2.1 Types of Eye Movements

There are a few basic types of eye movements that serve two purposes – to shift gaze direction to regions of interest, and maintain gaze at regions of interest. These eye movements are: saccades, smooth pursuit, vergence, vestibulo-ocular, and optokinetic movements (Purves et al., 2012). As mentioned previously, the purpose of making eye movements is to bring regions of interest to the fovea where visual acuity is the highest. It is also possible to position the region of interest on the fovea by moving the body or head

without moving the eyes' direction relative to the body/head, but this would be an inefficient way to sample a scene as it requires far more muscular effort.

Saccades are rapid, ballistic eye movements that are generally followed by a fixation, which is the time when the eyes are relatively stationary to maintain visual gaze on a single location to receive information (Gilchrist, 2011). They are considered ballistic because once a saccade landing position has been planned and initiated, the trajectory of the saccade cannot be modified. If the saccade missed the intended target, another corrective saccade will be made to bring the target to the fovea. Corrective saccades can be made if the target was missed due to an undershoot (Palmer, 1999), or if the target moved (Purves et al., 2012). Saccadic movements are also conjugate, which means that both eyes move simultaneously in the same direction with the same amplitude, which is the angular distance the eye travelled. However, if vergence movements are involved, which is the simultaneous movement of the eyes in opposite directions for foveation at targets that have a change in depth, the direction and amplitude of both eyes will not be the same (Land & Tatler, 2009).

Saccades have a characteristic temporal profile, where the velocity of the saccade increases until a peak velocity is reached, followed by rapid deceleration just before the saccade destination is reached (Land & Tatler, 2009). They are made up of extremely fast movements, with peak velocities reaching 900°/s (Palmer, 1999). However, the velocity and duration of a saccade depend on the amplitude of the saccade. There is a linear relationship between the duration of a saccade and the amplitude (Land & Tatler, 2009), and peak velocity and amplitude when plotted on log-log axis (Gilchrist, 2011). That is, the larger the saccade size (amplitude), the longer the saccade duration and the higher the peak velocity. Also, if one considers how fast saccades are, it is surprising that motion blur is not perceived. This is largely because during saccades, visual perception is selectively blocked so that the

motion blur and gap in perception is not perceived. This phenomenon is referred to as saccadic suppression (Palmer, 1999).

Smooth pursuit eye movements are a slower form of eye movements to maintain foveation on a moving target (Purves et al., 2012). This ability to retain foveation on a moving target is important because the visual resolution in the periphery is lower compared to the fovea. Additionally, if a target is moving at speeds greater then 15°/s, catch-up saccades will be made to bring the target back onto the fovea. Moreover, when a stationary target begins to move, there is a latency of 100-150ms before pursuit movements begin. To make up for the positional error caused by this delay, a catch-up saccade is initially made (Land & Tatler, 2009). Also, it is nearly impossible to make smooth pursuit movements in the absence of a moving target. Only highly trained observers can produce smooth pursuit movements without a target, while most individuals will produce rapid saccades instead (Purves et al., 2012).

Smooth pursuit eye movements are different from saccades in a number of ways. First, saccades are jerky and abrupt eye movements, while pursuit movements are smooth. Second, during eye movement, saccades do not rely on visual feedback, while smooth pursuit movements do. This is because the latter requires a constant update of signals between the brain and eye as the movement is being planned by the updated location of the target. Third, saccades are much faster than pursuit movements, which vary depending on the velocity of the target motion, but have a maximum velocity of 100°/s (Palmer, 1999). Beyond 100°/s, pursuit movements become purely saccadic (Land & Tatler, 2009). Finally, scene perception is not affected during saccades due to the lack of motion blur as a result of saccadic suppression, but for pursuit movements, the scene beyond the tracked target is blurry (Palmer, 1999).

The final type of eye movement that is used to alter gaze direction is vergence movement. When an individual fixates on a target, their eyes are positioned so that the target is aligned with each fovea. Therefore, to fixate on a near target, the eyes rotate towards each other (convergence), and to fixate on a far target, the eyes rotate away from each other (divergence) (Palmer, 1999). Vergence movements are smooth in nature, and have a latency of approximately 160ms (Land & Tatler, 2009). The velocity of vergence movements is also relatively slow, with peak velocities rarely surpassing 10°/s (Palmer, 1999). The primary difference between the other eye movements discussed and vergence is that vergence movements are disconjugate (i.e. move in different directions).

The three eye movements discussed thus far primarily operate while the head position is fixed. The next two eye movements that will be discussed – vestibulo-ocular & optokinetic movements – operate together to maintain foveation of a target while the head or body moves (Purves et al., 2012). These reflexive movements prevent the visual scene from shifting on the retina while head position moves. An example of vestibulo-ocular reflexive (VOR) movement is to fixate on an object, and move one's head from right to left. If, for example, an object being fixated was straight ahead, and the head was moved to the left, foveation will be maintained on the object by making a compensatory eye movement to the right (i.e. opposite direction). This eye movement is considered vestibular because it operates based on information from the vestibular system, which is the part of the human sensory system that controls balance. It is located in the middle ear where information about head position and orientation is sent to the brain (Palmer, 1999).

Optokinetic movements are eye movements made when large portions of the visual field are moved uniformly. It is a combination of fast and slow-phase movements. An example would be looking out the rear window of a moving car. It is characterised by slow eye movements in

the direction of the stimulus (in this case the road), and once the stimulus moves out of the field of vision, a rapid movement is made to return the eye position to the initial starting point. This alternating fast and slow-phase movements in response to the visual scene is known as optokinetic nystagmus, which is a normal reflexive movement of the visual system. This is not to be confused with pathological nystagmus, which is the result of an impaired visual system. Unlike vestibulo-ocular movement that operates based on information from the inner ear, optokinetic movements are driven by motion in the visual field (Purves et al., 2012). This may seem like smooth pursuit movement, but the primary difference between pursuit and optokinetic movement is that the latter operates on a large portion of the visual field, while the former on small target only (Land & Tatler, 2009).

Both the vestibulo-ocular and optokinetic movements act in a complementary manner due to their dependency on the speed of motion of the visual field. The vestibulo-ocular response is comparatively insensitive to slow movements, unlike the optokinetic response, which is sensitive to slow movements. As such, at high frequencies, the vestibulo-ocular response augments the optokinetic response, and vice versa for low frequency movements (Purves et al., 2012). A primary difference between both movements is that the optokinetic response is part of a feedback system, while the vestibulo-ocular response is not. The optokinetic response operates by matching the velocity of the visual scene to the rotation and movement of the eye, thus cancelling out the displacement of the scene. The vestibulo-ocular response does not rely on a feedback system because movements of the eyes do not affect the vestibular system (Land & Tatler, 2009). Finally, both the vestibulo-ocular and optokinetic movements are much more rapid and accurate compared to smooth pursuit movements (Palmer, 1999). The latency between head movement and the resulting eye movement of the vestibulo-ocular response is no greater than 15ms (Land & Tatler, 2009).

Lastly, there are two types of small eye movements that are constantly being made – ocular-drifts and micro-saccades. Ocular-drifts are small, slow movements of the eyes that shift fixation away from the target. They are generally corrected with micro-saccades, which are similar to regular saccades, except micro-saccades have smaller amplitudes – between 2 and 120 arcminutes (Rayner, 1998). These micro eye movements are made to prevent the visual image from fading due to neural adaptation (Gilchrist, 2011). Thus, when fixations are referred to as periods of time when the eyes are stationary, that is not completely accurate as minuscule movements are constantly being made.

7.2.2 Saccade Latency and Amplitude

In most experimental work, saccade latency is the time taken to initiate a saccade after target onset. However, when a saccade is part of a string of saccades (scan path), it corresponds to the inter-saccadic interval. One of the most prominent characteristics of saccade latency is the distribution of latencies. Regardless of the task, stimuli, or expertise, the distribution of saccadic latencies takes the form of an asymmetrical distribution with a positive skew for longer latencies (Sumner, 2011). These latencies can range from 100 - 1000ms (Gilchrist, 2011), with the majority of saccades occurring between 150 - 200ms for externally triggered saccades (Land & Tatler, 2009).

Saccade latencies can be affected by several factors: 1) Exogenous saccades are faster than endogenous saccades. Exogenous saccades are triggered by stimulus or target onset, while endogenous saccades are goal-driven and occur in the absence of a specific target at saccade destination (Sumner, 2011); 2) Targets with greater intensity/saliency will elicit shorter latencies (Sumner, 2011; Gilchrist, 2011); 3) Targets further away from the fovea have longer latencies (Land & Tatler, 2009; Gilchrist, 2011), with shorter latencies being elicited when target eccentricities are between 2 - 10° (Sumner, 2011); 4) Temporal and spatial expectancy

can also reduce saccade latencies (Sumner, 2011); 5) The speed-accuracy trade-off typically observed in many tasks is exhibited here – where fast latencies are more likely to be erroneous in saccade landing position, either in a completely incorrect direction or an under/overshoot of saccade destination (Sumner, 2011; Rayner, 1998); 6) When saccades are made in the context of many possible locations, latencies are greater with more potential positions (Sumner, 2011); 7) Introducing a gap of about 200ms between fixation offset and target onset (gap paradigm) will decrease saccade latencies greatly (Sumner, 2011; Gilchrist, 2011; Rayner, 1998); 8) When a target is presented simultaneously with a distractor, saccade latencies will be greater, and the landing position of the saccade will be to an intermediate location between target and distractor (Sumner, 2011; Gilchrist, 2011; Rayner, 1998); 9) Cognitive process involved can also influence saccade latency. When several saccades are planned ahead, latencies are greater. Moreover, in the context of an anti-saccade paradigm (saccade to be directed away from target), latencies are also greater (Rayner, 1998; Gilchrist, 2011). However, while the factors above may influence the overall average saccade latency, the shape of the distribution will always remain the same (positive skew).

The distributions exhibited by saccade latencies are typically unimodal. However, there is a separate category of extremely fast saccades, called express saccades (with latencies <100ms), which exhibits a pattern of bimodal distribution of saccade latencies. The possibility of evoking express saccades depends on the experimental conditions as mentioned above. Fischer and Boch (1983) found that monkeys were able to elicit express saccades when a gap with no visible stimulus was introduced between fixation offset and stimulus presentation. With gap durations of 200-240ms, saccade latencies were about 70-80ms, with low standard deviations (± 3ms). There was a negative linear relationship between gap durations and saccade latencies (i.e. as gap durations decreased, saccade latencies increased).

Next, we discuss why saccade latencies vary on a trial-by-trial basis when the task and stimuli are kept constant. The initiation of a saccade is believed to occur once the discharge rate of the neuron(s) responsible for encoding a saccade reaches a threshold level of neuronal activity. Several models have been developed to demonstrate the latency distributions observed in saccades (e.g. Carpenter & Williams, 1995). The common framework of these models depicts a decision or motor unit that rises from a starting level, S, until a threshold, T, is reached, which then initiates the saccade (Sumner, 2011). According to the model of Carpenter and Williams (1995), the starting level (S) of activity is not always constant. Instead, they propose that the starting level of activity will be varied depending on the prior probability of a target being present. For instance, if there is a high probability of a target appearing at the same location each time, the decision signal (starting level) will be increased (S_1) compared to a condition where the possible location of the target is unknown (S_0). Therefore, the activity needed to reach the threshold is lower, which in turn produces a saccade with short latency. Furthermore, the model posits that the rate of rise-to-threshold, while always linear, varies between trials. As such, due to the variations in the rise-to-threshold, variations in saccade latencies are also produced. A possible source of variance for the rise-to-threshold is perceptual processing (Sumner, 2011), where the brain requires time to make a decision about the presence of a target. And since the decision making process is stochastic in nature, it could reflect the distribution of latencies observed.

Saccade amplitudes also vary depending on the type of tasks being performed. The average saccade amplitude for silent reading is 2°, while it is 1.5° for oral reading, and 1° for music reading. The saccade amplitudes are larger for visual search tasks (3°) and scene perception (4°) (Rayner, 1998). However, von Wartburg et al. (2007) found that saccade amplitudes (as measured by mean and median values) to scene images scaled proportionally with image size. Which implies that our saccadic system has a scaling ability. Furthermore, saccades

made during active tasks (like tea making (Land & Hayhoe, 2001)) have much greater amplitudes then sedentary tasks e.g. reading (Land & Tatler, 2009). Land and Hayhoe (2001) found that saccade amplitudes for within-action saccades were typically between 5° and 10°, with the task being performed not influencing the saccades. Conversely, between-action saccades (i.e. first saccade to a new object) varied with the task, whereby a sandwich making task (on a countertop) had saccade amplitudes between 5° and 30°, and a tea making task (in a larger kitchen area) had amplitudes ranging from 10° to 90°.

Finally, it is important to note that saccades do not always land precisely on the target destination. Especially for targets that are far from the fovea, there tends to be a 10% undershoot to the target, followed by a corrective-saccade that will make up for the error after about 100-150ms (Land & Tatler, 2009). However, those findings are characteristics of laboratory tasks, which was not found by Land and Hayhoe (2001) in active tasks as participants were able to make accurate saccades to the object they were interested in.

7.2.3 Eye Movement Control

As stated before, eye movements are made with the purpose of positioning or maintaining areas of interest on the fovea. However, that raises the question of what aspect of a scene induces a saccade and subsequent fixation. This section will first briefly describe the underlying physiology of oculomotor control. Following this, the two schools of thought – bottom-up & top-down – regarding the mechanisms that drive eye movements to certain locations of a scene are discussed. Bottom-up explanations of eye movements are centred on the notion that eye movements are driven by properties of a scene that is on the retina. In other words, properties of a scene that are salient (being noticeable) will be looked at, while not taking into account any contributions from higher cognitive processes. Top-down

explanations state that eye movements are goal-directed, thereby involving high-level cognitive control (Land & Tatler, 2009).

Neural control of voluntary saccades can be traced to the Frontal Eye Fields (FEF), which project to two efferent pathways. The first projection is to the Superior Colliculus (SC), whose function is to encode the position of the planned gaze direction. The resulting output from the FEF and SC projects to two premotor nuclei situated in the midbrain – Paramedian Reticular Formation (PRF) and Rostral Interstitial Medial Longitudinal Fasciculus (riMLF). The PRF and riMLF control the amplitude and velocity of the horizontal and vertical components of the saccade, respectively. The neural pathway of VOR begins with the semicircular canals – horizontal, superior, and posterior canals – that transform head rotation to neural signals. These canals are opponent pairs, which means that activity on one side of the head inhibits activity of the contralateral pair on the opposite side of the head. The vestibular nerve carries the neural signal from the semicircular canals to the vestibular nuclei, which subsequently provide information about eye velocity signals to the oculomotor neurons. For vergence movements, the signal originates from primary visual cortex (V1), which can detect binocular disparity and blur. Vergence movements are then initiated by the FEF that signals the premotor Nucleus Reticularis Tegmenti Pontis (NRTP) in the midbrain (Schor, 2010).

Cortical motion signals from the Medial Temporal (MT) cortex, and the Medial Superior Temporal (MST) cortex control smooth pursuit movements due to the need for constant visual feedback. The MT receives inputs from the primary visual cortex and encodes a three-dimensional representation of the speed and direction of the target relative to the eye, which is then projected to the MST and FEF. From here, projections are sent to the Dorsal Lateral Pontine Premotor Nuclei (DLPN) to generate smooth pursuit movements. The DLPN plays a

major role in preserving the steady-state velocity of the smooth pursuit movement. To generate optokinetic nystagmus (instead of smooth pursuit), MST and FEF outputs have to be transmitted ipsilaterally (same side) to the nucleus of the optic tract. The MST is crucial for both optokinetic nystagmus and smooth pursuit as it the cortical region that coordinates motion signal. Visual stimuli evoke optokinetic responses by stimulating ganglion cells in the retina that respond solely to motion in specific directions or orientations. This information passes through the optic nerve, crosses at the chiasm, and is transmitted to either the cortex or the midbrain (via the lateral geniculate nucleus and optic tract, respectively) (Schor, 2010).

Bottom-up explanations of eye movements are salience-driven. An example of a salience-based model was developed by Itti and Koch (2000) (see Figure 7.1 for diagram of model). The model begins by first extracting in parallel low-level features of the entire scene. This mimics the extraction of visual information by the early stages of the visual system (V1), which includes features such as orientation, colour, and brightness at different spatial scales. Each of the features that are processed in parallel produce a feature map, which are then combined to produce a single salience map. This salience map will have several hot spots that correspond to areas of interest in a scene. A "winner-takes-all" mechanism operates to ensure only one part of a scene is highly salient, and this will coincide with the destination of the saccade. The model also incorporates an inhibitory mechanism that suppresses the last attended location, which allows the model to move on to a new location that is lower on the saliency map. As such, the model does not get stuck on a single most salient location, but moves down in a hierarchical manner to all other hot spots. This concept of a salience map was initially introduced by Koch and Ullman (1985) to achieve pre-attentive selection.

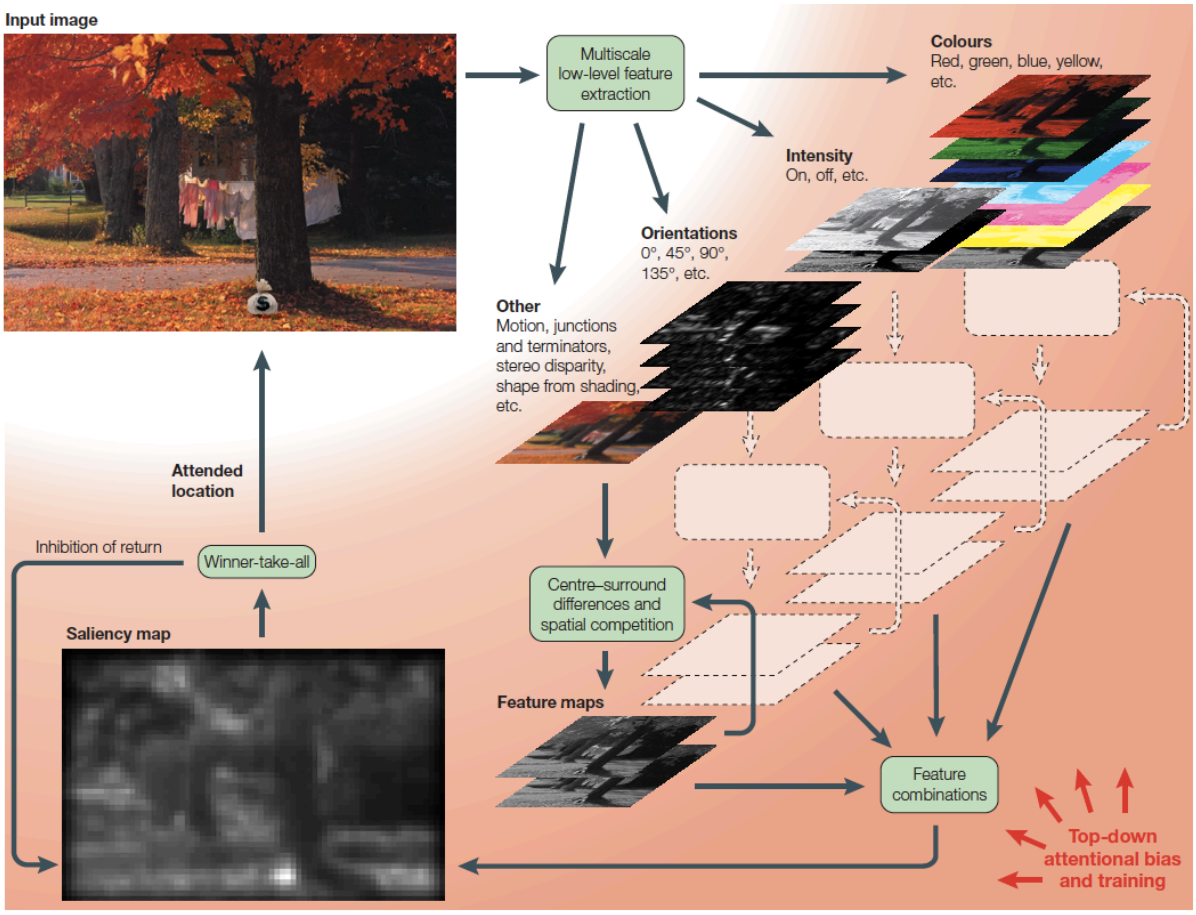


Figure 7.1. Diagram of Itti and Koch's model for the control of bottom-up attention. From Itti and Koch (2001).

Some studies (e.g. Parkhurst, Law, & Neibur, 2002; Parkhurst & Neibur, 2003; 2004) have shown evidence of a salience-based mechanism operating when viewing natural scenes. Parkhurst et al., (2003) found that bottom-up mechanisms contribute to the guidance of eye movements in free-viewing conditions. Participants' eye movements were recorded while they free-viewed stimuli of natural (home interiors, natural landscapes, and buildings/city scenes) or artificial (fractal images) scenes. Each stimulus was displayed for five seconds, and participants were instructed to "look around at the images". Using a bottom-up salience model (Koch & Ullman, 1985), the experimenters calculated the average salience at the point of first fixation made by the participants. They found that the mean salience was significantly higher than the mean salience expected by chance (the average salience of randomly chosen

points on the stimulus). They took this result to indicate that the location of first fixation was guided by the salient properties of the stimuli. Moreover, they found that the mean salience of subsequent fixations was progressively lower, but still above the levels expected by chance.

Parkhurst, Law, and Neibur (2003) further analysed the relative contributions of the different features (colour, orientation, and luminance). That is, instead of the salience being calculated from the combination of the three features, each sub-modality was calculated individually. The authors found that overall, the colour and luminance feature channels contributed more than orientation feature. However, a notable exception to this is for the building/city scene category, where the orientation feature contributed more than the colour feature. The colour feature contributed most for the fractals and home interior scenes, while the luminance contributed most for landscapes and building/city scene. This suggests that the relative contribution of the different features depends on the types of scenes viewed.

However, a salience-based model of eye movement control does not take into account several other aspects of eye movement control. Namely, the predictive power of salience-based features is lower than the predictive power of object-based information. Furthermore, behavioural biases such as preferring to look at the centre of the screen, making small amplitude saccades instead of larger amplitudes, and making more horizontal movement compared to vertical or oblique movements could also play a role in eye movement control. Finally, while most studies do show a difference in image properties between fixated and control locations, the magnitude of differences is quite small (Land & Tatler, 2009).

Further limitations to this account of eye movement control are in the design of the model itself. First, an inhibiting mechanism, such as the inhibition of return (IOR) (e.g. Posner & Cohen, 1984; Klein, 2000), will only be temporary, as the effect will eventually wear off, after which eye movements should once again return to a location that has already been

viewed. As such, patterns of fixation will show a repeating cycle of fixations. Another limitation is that the saliency map is constructed in parallel across the entire visual scene. However, the resolution across the retina changes, which means that a saliency map cannot be constructed all at once in parallel because eye movements with large amplitudes cannot be targeted due to information that is not detectable further in the periphery.

As outlined above, a bottom-up salience account of eye movement control cannot fully explain the patterns of fixation we see in everyday life. Therefore, it is plausible that top-down knowledge is also influencing eye movement control. This is evident in studies that demonstrate how eye movements are influenced by higher-level control of task demands. Yarbus (1967) found that participants had different patterns of eye movements depending on the type of question the experimenter posed. Figure 7.2 shows the same participant's eye movements to the picture 'unexpected visitor' under various conditions. Each question (Figure 7.2c-h) resulted in different patterns of fixations, which were also markedly different from the free-viewing (Figure 7.2b) condition. It is also important to note that studies regarding the saliency of a scene are always conducted under free-viewing conditions to avoid 'cognitive demands'. But in real-world settings, are our minds really blank while we look around? For the most part, our vision is utilized to guide actions to match our needs.

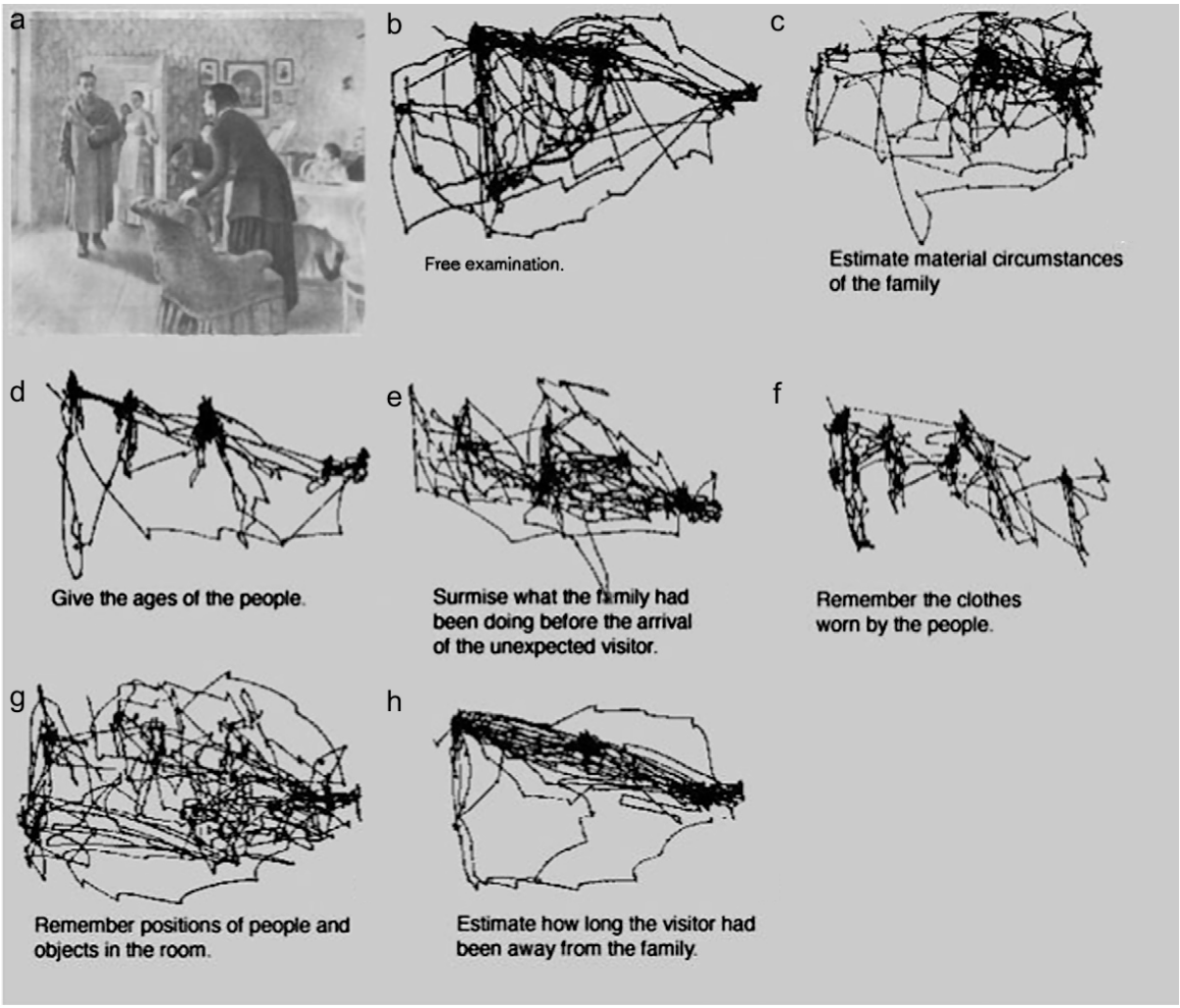


Figure 7.2. Example of one participant's eye movements during three minute recordings of a scene titled 'unexpected visitor' (a). (c-h) Different patterns of fixations are evoked by different questions raised. From Yarbus (1967).

Castelhano, Mack, and Henderson (2009) also studied how task (visual search for an object & memorization) influenced eye movements. They found that the type of task influenced eye movements in several aspects such as the number of fixations and areas fixated on. For both of these measures, the memorization task had a greater number of fixations and areas fixated on compared to the visual search task. However, average saccade amplitudes and fixation duration were the same for both tasks. This is especially interesting because according to the literature on eye movement control during reading (Rayner, 1998), fixation duration is influenced by the task, context, word length, and difficulty. This implies that the manner in

which high-level cognitive demand influences eye patterns fixations is very varied. Regardless, high-level control does influence eye movements, which challenges the bottom-up salience approach.

The literature discussed so far has involved studying eye movements to scenes/natural images. However, our interest lies in the effects of texture properties (specifically the orientation domain) on eye movements. This will be discussed in detail, in Section 7.5.

7.3 Measuring Eye Movements

The process of measuring a point of gaze, whether stationary or in motion, is referred to as eye tracking, and therefore the devices used to measure these eye movements are called eye trackers. They are important not just in the field of research, but also for commercial use. Studying eye movements allows us to gain insight in a number of ways. By knowing what an observer is looking at, it is possible to infer what in a visual scene they find interesting, and even attempt to understand how perception occurs (Duchowski, 2007). This has been a great source of interest for over a century, with Dodge and Cline (1901) reporting one of the first pieces of photographic documentation of eye movement recordings.

From then on until the late 1950s, eye movement studies were mostly conducted in laboratory settings with bench-mounted devices that recorded eye movements of restrained observers. A key outcome from those studies was that fixations tend to lead actions (i.e. fixations are made to objects a few seconds before participants make contact with the object) (Land and Tatler, 2009). Many fundamental facts about eye movements such as saccadic suppression, saccadic latency, and the size of perceptual span were also discovered during this period (Rayner, 1998). Norman Mackworth advanced eye movement studies in the 50s by creating the first mobile eye tracker (Land and Tatler, 2009). This device consisted of a small movie camera

and a periscope that transmits a spot of light reflected by the eye. Within the camera, a spot that marks the observer's line of sight is immediately marked onto the image of the scene. Therefore, as the observer moves about a scene, the device records the dynamic scene while taking into account eye movement (Mackworth & Thomas, 1962). However, these devices were still quite bulky, and only in the 1990s did eye movement studies become realistic to conduct outside the laboratory. This was mostly due to the invention of smaller cameras, and computers that could assist in data analysis (Land & Tatler, 2009).

There are four broad categories of eye movement measurements – Electo-OculoGraphy (EOG), Scleral Search Coil, Photo/Video-OculoGraphy (P/VOG), and Video-based Combined Pupil/Corneal Reflection. The EOG is an invasive method of eye tracking, as it requires electrodes being placed around an observer's eyes. The position of the eye can be estimated based on the different electric potentials, which are in the range of 15-200μV. The nominal sensitivity (the 'natural' sensitivity of the sensor) for the device is 20μV/degree of eye movement. For this method of measurement, it is necessary to know the position of the head, as eye movements are relative to the head position (Duchowski, 2007). A major advantage of this method is the ability of the EOG to record horizontal and vertical eye movements separately, and also to record eye movements while the eyes are closed (Chennamma & Yuan, 2013).

The Scleral Search Coil is another extremely invasive method of eye tracking, but with very precise measurement that is accurate up to 5-10 arc-seconds over a range of 5°. This method involves the use of a large contact lens (that covers both cornea and sclera) that has a coil of wire attached to it. While wire coils are the most common attachment to the lens, reflecting phosphors and line diagrams have also been used. With the wire coil, eye position is measured as the wire coil moves through an electromagnetic field (which produces a voltage

in the coil that can be measured). Like the EOG, this method requires the known position of the head (Duchowski, 2007).

P/VOG is a form of eye tracking that can be both invasive (head-mounted) and non-invasive (remote) (Chennamma & Yuan, 2013). This method groups together a wide variety of measurements that are able to distinguish features of the eyes such as pupil shape, limbus position (boundary between iris and sclera), and corneal reflection. Like the other two methods, this method requires the position of the head to be known, but most often the head position is immobilized with the use of a head/chin-rest or bite-board. The P/VOG method typically relies on offline frame-by-frame analysis to measure the position of the eye relative to the head. This form of eye tracking (as well as the EOG and Scleral Search Coil) does not readily provide point of regard (POR) measurements, which is a real-time estimation of a participant's gaze (x-y coordinates) on the visual display (Duchowski, 2007).

The final method of Video-based Combined Pupil/Corneal Reflection eye tracking is distinguished from the other forms of eye tracking by its ability to measure POR i.e. measure the orientation of the eye in space, as opposed to measuring the position of the eye relative to the head. This is primarily because this method uses multiple ocular features (corneal reflection and pupil centre) to separate eye movements from head movement. The process involves the use of a light source (usually infrared or near-infrared) to reflect on the cornea and pupil of the observer. The relative positions of the pupil centre and corneal reflection are typically used to measure the POR, which is then mapped onto the visual scene (Duchowski, 2007).

The corneal and other reflections, also known as the Purkinje reflections, are a rich source of information due to the four reflections formed (Figure 7.3a). The first Purkinje image is formed by the reflection of light from the front of the cornea. The second image is formed by

the reflection from the rear of the cornea. These two reflections almost coincide with each other. The third Purkinje image is formed when light that passes through the cornea and aqueous humor to reflect off the front surface of the lens. Finally, the fourth Purkinje image is formed by the reflection of light from the rear end of the lens. Should the eye move due to lateral head movements (side-to-side), the first and fourth Purkinje images will move the same distance. However, if the eye rotates, the first and fourth Purkinje image will move at different distances (changing their separation). This separation between the first and fourth Purkinje images serves as a measure of the angular orientation of the eye, which is insensitive to head movement (Crane, 1994).

While some eye trackers use dual-Purkinje images (first and fourth reflection,) the first Purkinje reflection is typically used in conjunction with the pupil centre to measure a participant's POR. Figure 7.3b shows how the position of the first Purkinje image relative to the pupil centre changes as an eye rotates to fixate on 9 locations on a calibration screen. Due to the infra-red source being placed at a fixed position, the first Purkinje image (small white circle) is stable while the eye moves, and the pupil on the other hand (black circle), changes position with the rotation of the eyes. Therefore, with minor head movements, the positional difference between the first Purkinje image and the pupil centre remain rather constant. However the positional difference between the two changes with eye rotation, allowing a POR measurement (Duchowski, 2007).

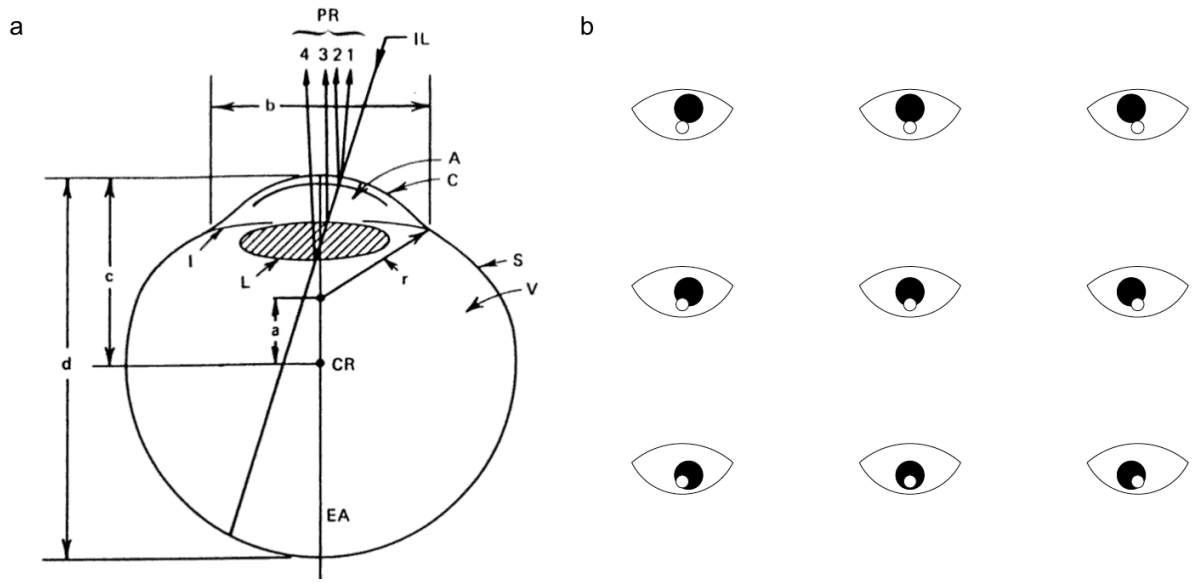


Figure 7.3. a) The Four Purkinje images formed (PR). IL=incoming light; A=Aqueous Humor; C=Cornea; S=Sclera; V=vitreous humor; I=iris; L=lens; CR=centre of rotation; EA=eye axis; a~6mm; b~12.5mm; c~13.5mm;d~24mm;r~7.8mm. b) the relative positions of the first Purkinje image and pupil. From Duchowski (2007).

The remainder of this section will be focusing on the eye tracker developed by Tobii, a Swedish based company. Tobii eye trackers use the method of Pupil Centre Corneal Reflection (PCCR), which is a form of the Video-based Combined Pupil/Corneal Reflection method. The remote and non-intrusive method is the screen-based eye trackers. These trackers are table mounted and have a similar appearance to any flat panel displays, but with the inclusion of projectors, sensors, and image processing algorithms embedded within the display. The essential components of these eye trackers are the illuminators (near-infrared light source), cameras, and processing unit. In the Tobii eye trackers, the processing unit contains algorithms for image detection, 3D eye modeling, and gaze mapping.

The process involved in the Tobii eye tracking begins with the illuminators. They create a pattern of near-infrared illumination on the eyes that is invisible to the naked eye, but visible to the two sensors, which take high-resolution images of the observer's eyes and their respective reflection patterns. This data is sent to the Tobii Eye Tracker (TET) server, where

image processing algorithms and a physiological 3D model of the eye interpret the reflection patterns from the sensors to generate the eyes' position and gaze point. Tobii uses two sources of ocular information to determine gaze direction – pupil reflection and corneal reflection (first Purkinje image) (Tobii Technology, n.d. a). As explained earlier, this is a robust method to calculate gaze position as the relative position between pupil centre and the first Purkinje image will remain relatively constant with minor head movements.

With the PCCR method used by Tobii, there can be two possible illumination setups – bright and dark pupil tracking. The primary difference between the two is the location of the illuminator with respect to the optical axis of the imaging device (Figure 7.4). When the light source is aligned with the imaging device, the images captured will show that the pupil appears brighter than the iris, which is due to the retroreflection causing the bright pupil effect. However, if the light source is placed away from the optical axis, the pupil will appear darker then the iris, thus causing the dark pupil effect. The bright pupil method is susceptible to external factors such as age and environmental light as both can have an effect on pupil size, which influences the ability to track the eye. Furthermore, both the dark and light pupil methods are sensitive to iris pigmentation. The bright pupil method is more ideal for Caucasians and Hispanics, while the dark pupil method is more ideal for Asians. It is therefore impossible to say that one method is definitely better than the other. As such, Tobii uses both dark and bright pupil to track gaze direction (Tobii Technology, n.d. b). During the calibration process, the observer will have both methods administered, and the method that yielded the highest accuracy will be used for the recording. However, if the data received during the recording shows an inability to track the eye, the method can change throughout the recording until the data is once more reliable (Tobii Technology, n.d. c).

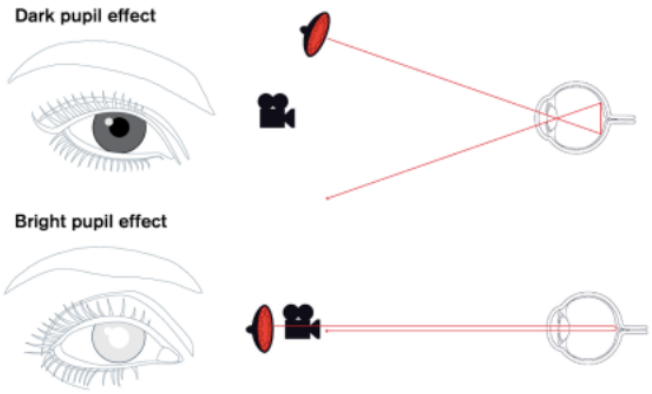


Figure 7.4. Example of dark (top) and bright (bottom) pupil method. From Tobii Technology (n.d. b).

Before the eye tracker can be used on an observer, the observer has to go through a calibration process. The purpose of the calibration is to optimise the accuracy of the eye tracker by measuring the characteristics of the eyes (how the eyes reflect light at different areas of the screen, Figure 7.3b). This process requires the observer to look at specific points (5 or 9 calibration points) on the screen. During this process, several images will be captured and analysed, with the resulting output being integrated with the physiological 3D model of the eye to produce a gaze point for each image captured (Tobii Technology, n.d. c). After the calibration, the accuracy of the calibration is shown for all the calibration points. Experimenters have the option to recalibrate points that are not very accurate (no gaze data obtained, or gaze data with large offset to calibration point).

Accuracy of an eye tracker is the average difference between the actual gaze position and the measured gaze position, while precision is the ability of the device to reliably produce the same measured gaze position over different trials. The Root Mean Square was used to measure the variation in data (Tobii Technology, 2011). Accuracy of the data collected is important for validity of the results, and this is illustrated in Figure 7.5. Weigle and Banks (2008) conducted a study to test the accuracy of a remote Tobii eye tracker (T60 model). The

test stimulus used was of a ball (50 pixel diameter) bouncing across several squares. They had one participant track the motion of the ball, either at the top or bottom edge of the ball in separate sessions. They found that the latency between the display of video frame (actual target position) and the resulting image capture (gaze report) of the eye was less than 1ms. Additionally, they computed the cross-correlation, which is the measure of similarity between two data sets as a function of the displacement from each other, between the coordinates of the actual target position and the corresponding gaze reports.

For accuracy of gaze data, visual inspection (of overlaid gaze report data on a time-lapse image of the ball bouncing) showed that eye tracker was able to track the motion of the eye as it followed the ball. However, it was found that the location of the eye-gaze position was not reported accurately. Weigle and Banks (2008) compared the difference between gaze location and target position for both the top and bottom-tracking of the ball. The difference between eye-gaze and actual target position was 4.4 pixels (x-coordinate) and 0.68 pixels (y-coordinate) when the participant was tracking the bottom edge of the ball. When the participant was tracking the top edge of the ball, the difference in eye-gaze and target position was 3.7 pixels (x-coordinate) and 34.2 pixels (y-coordinate). An important aspect of this study was that the experimenters knew beforehand the intended gaze location (the top and bottom edge of the ball). If this were not the case, the data would show that the participant was tracking the centre of the ball instead of the bottom edge (as the accuracy was off by 34 pixels for the y-coordinate of the bottom tracking motion). Overall, this study suggests that accuracy can be tricky to ascertain, and that the Tobii eye tracker, even with calibration, is not always accurate. A few things to note about this study are that there was only one observer, and said observer did have difficulties with the calibration process. Also, multiple stimuli were used during testing, but they only reported the results of one stimulus.

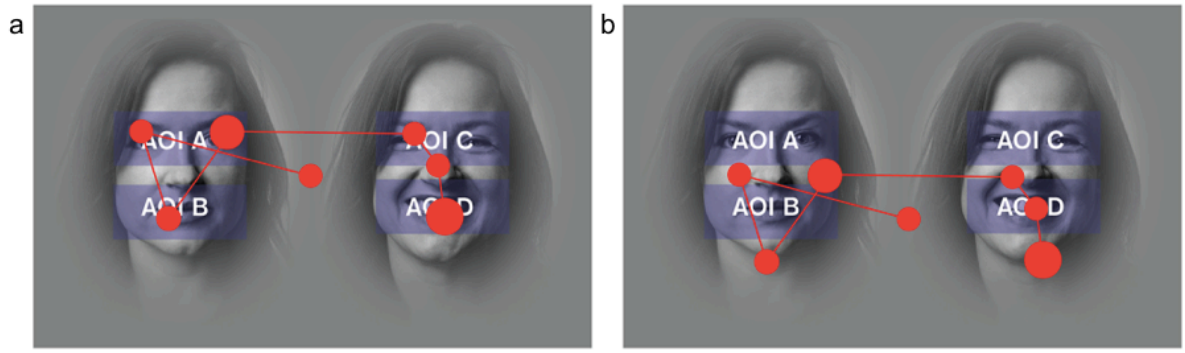


Figure 7.5. Example of gaze data with a) high accuracy and b) low accuracy. 6 out of the 7 fixations land on an Area of Interest (AOI) when accuracy is high, but when accuracy is low due to a vertical offset, only 1 fixation lands on an AOI (b). From Tobii Technology (n.d. d).

The Tobii eye tracker also reports a certain degree of latency, which is the time between an image capture (by camera sensors) and gaze data that is generated by the TET servers to an external application (e.g. Tobii Studio). This latency period is due to the actual time of image capture, transfer of information, and internal calculations. For the Tobii T60, this latency is around 30±3ms (Tobii Technology, 2010a). Tobii studio (described below) also reports a lag in the presentation duration of a stimulus. This offset is the difference in time between the value set in the properties box of the experimental setup and the actual duration of the stimulus during presentation. According to tests conducted, the offset is typically around 100ms for durations below 1000ms, and 110-120ms for durations above 1000ms (Tobii Technology, 2012).

The Tobii eye tracker that will be used for the current studies is the Tobii T60, which is a non-intrusive remote system. The eye tracker has been integrated into a 17-inch thin-film-transistor (TFT) monitor with a screen resolution of 1280×1024 pixels. The T60 samples data at 60Hz (16.67ms), and has an average accuracy of 0.5° with an average drift of 0.1°. The sampling rate refers to the number of gaze points sampled per second. Accuracy refers to the difference between the actual and reported gaze direction. Drift refers to the alteration in

accuracy as a result of a change in lighting. The spatial resolution (frame-by-frame variation of the measured gaze point) of the Tobii T60 is approximately 0.2°. Tobii technical specifications state that these measurements were obtained as an average response when tested in a "controlled office environment" with an observer situated 63cm away from the screen. This model allows for binocular tracking, using both dark and bright pupil tracking for optimization. At a viewing distance of 70cm, it allows for freedom of head movements up to 44×22×30cm (Tobii Technology, 2010b). Tobii Studio software will be used in conjunction with the eye tracker to design the experiment, as well as record and analyse of data. Tobii Studio has some built-in tools that can be used to visualise and analyse the data (e.g. heat maps, fixation count, fixation duration, etc.), or produce additional outputs (e.g. x-y coordinates of gaze point, pupil size, keyboard responses, etc.) (Tobii Technology, 2012).

7.4 Eye Movements to Non-Textured Targets

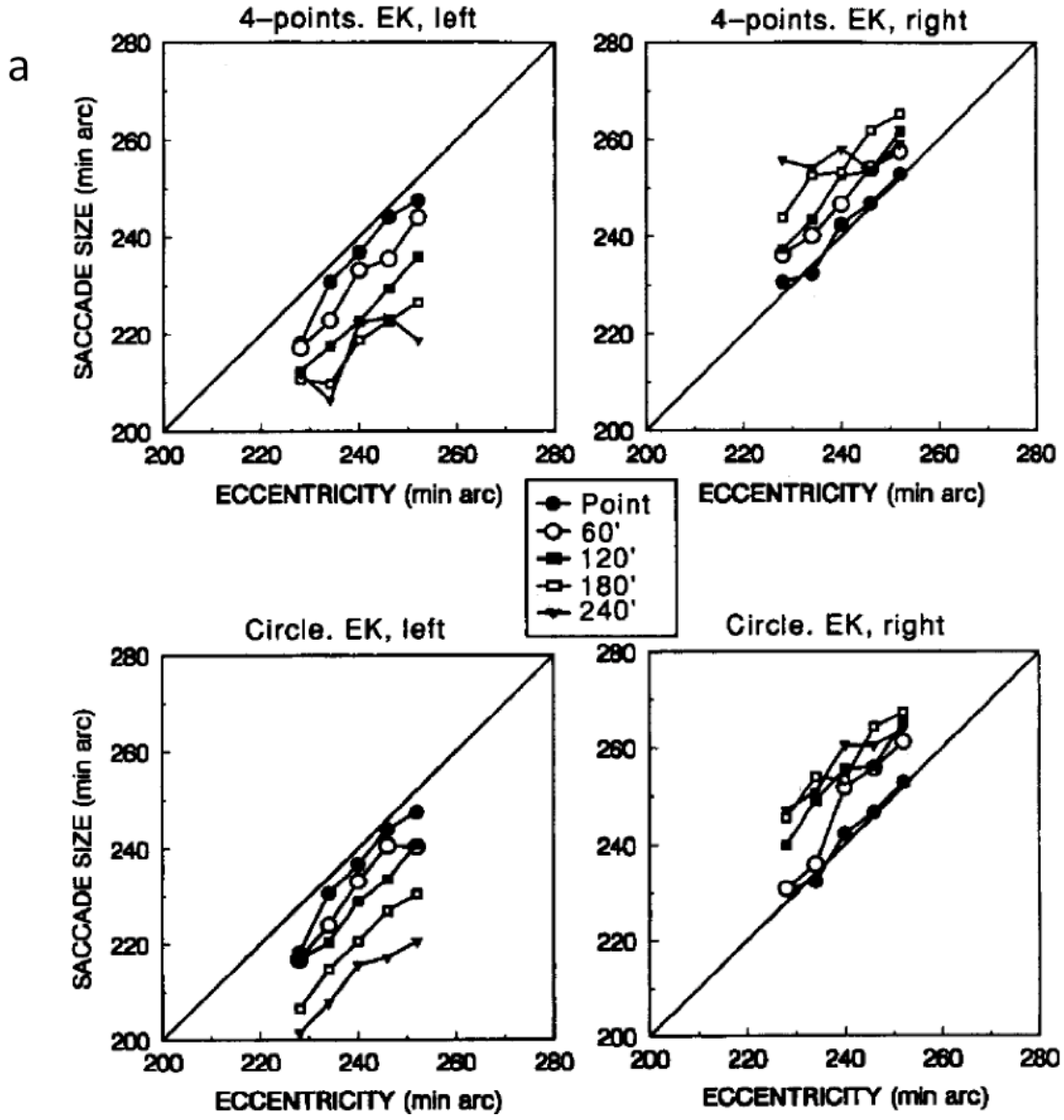
Many studies investigate eye movements in reading, visual search, and scene perception (e.g. see Rayner, 2009; Liversedge & Findlay, 2000 for reviews). Additionally, in the field of oculomotor research, eye movements to solitary targets (i.e. not embedded in a background) are also frequently studied. In this section, we briefly discuss the relevant studies that investigate the spatial and temporal aspects of eye movements to targets. The initial few studies that will be discussed focus on simple targets (e.g. circle or square shaped targets), while the last study investigates more complex shapes. The reader is directed to Table 7.1 and 7.2 for further details about the saccade amplitudes and standard deviations for studies discussed in this section.

As previously stated, saccades typically undershoot the intended target by approximately 10%. However, Kapoula and Robinson (1986) found a range effect of saccade accuracy – whereby saccades overshoot near targets and undershoot far targets. Their participants were

required to perform two tasks, which were 1) to make a saccade to a target that was a red spot subtending 5.6 arcminutes, and 2) to make a saccade to a target that could either be a normal E or backwards E (Ǝ) that was 24×12 arcminutes in size, and press a button when the target was a normal E (Visual Discrimination). The targets for both tasks could be positioned to right or left of fixation at eccentricities of 5°, 10°, 15°, and 20°. The different eccentricities were randomized within each block. The results were that for targets appearing at 5° eccentricity, there was a tendency to overshoot the target, and this was more prominent in the task where they had to discriminate the target compared to when they only had to make a saccade to the target. For targets at the remaining eccentricities, participants always undershot the target location. Interestingly, when 5° targets were presented alone (i.e. no other eccentricities tested within the block), the overshoot was no longer observed, with saccades once again demonstrating an undershoot to the target. The authors therefore refute the claim that saccadic undershoot always occurs. According to them, instances in which targets appear at a range of different eccentricities produce saccades that overshoot targets at small eccentricities and only undershoot targets at far eccentricities.

Kowler and Blaser (1995) also studied the accuracy of saccades to small and larger targets. The targets were either a single point, circle, or diamond. The latter two had 4 different diameters each – 60, 120, 180, and 240 arcminutes. The target eccentricities were between 228 and 252 arcminutes, with 6 arcminute increments. The target type, eccentricity, and direction (left or right) were randomized within each experimental session, but participants were always informed about the direction of the target beforehand. The participants were told to make an eye movement to the target with only 1 saccade, and were encouraged to increase saccade latencies for maximum accuracy. The results show that saccades were highly accurate (≈1.25% undershoot) for the single point targets at all eccentricities. For the circle and diamond targets, deviation from the centre of the target increased with the larger target

sizes. Meanwhile, saccade size increased proportionally to the eccentricity of the target (i.e. the average landing position with respect to the contour of the target remained the same). Saccadic precision, as measured by the standard deviation (SD) of saccade sizes, was between 12 and 16 arcminutes for single point targets, and targets with diameters up to 120 arcminutes. However, as target size increased beyond that, saccadic precision decreased (see Figure 7.6b). Finally, they found that saccade latencies did not change with the target size or eccentricity. On average, saccade latency was 240ms and 320ms for the two participants.



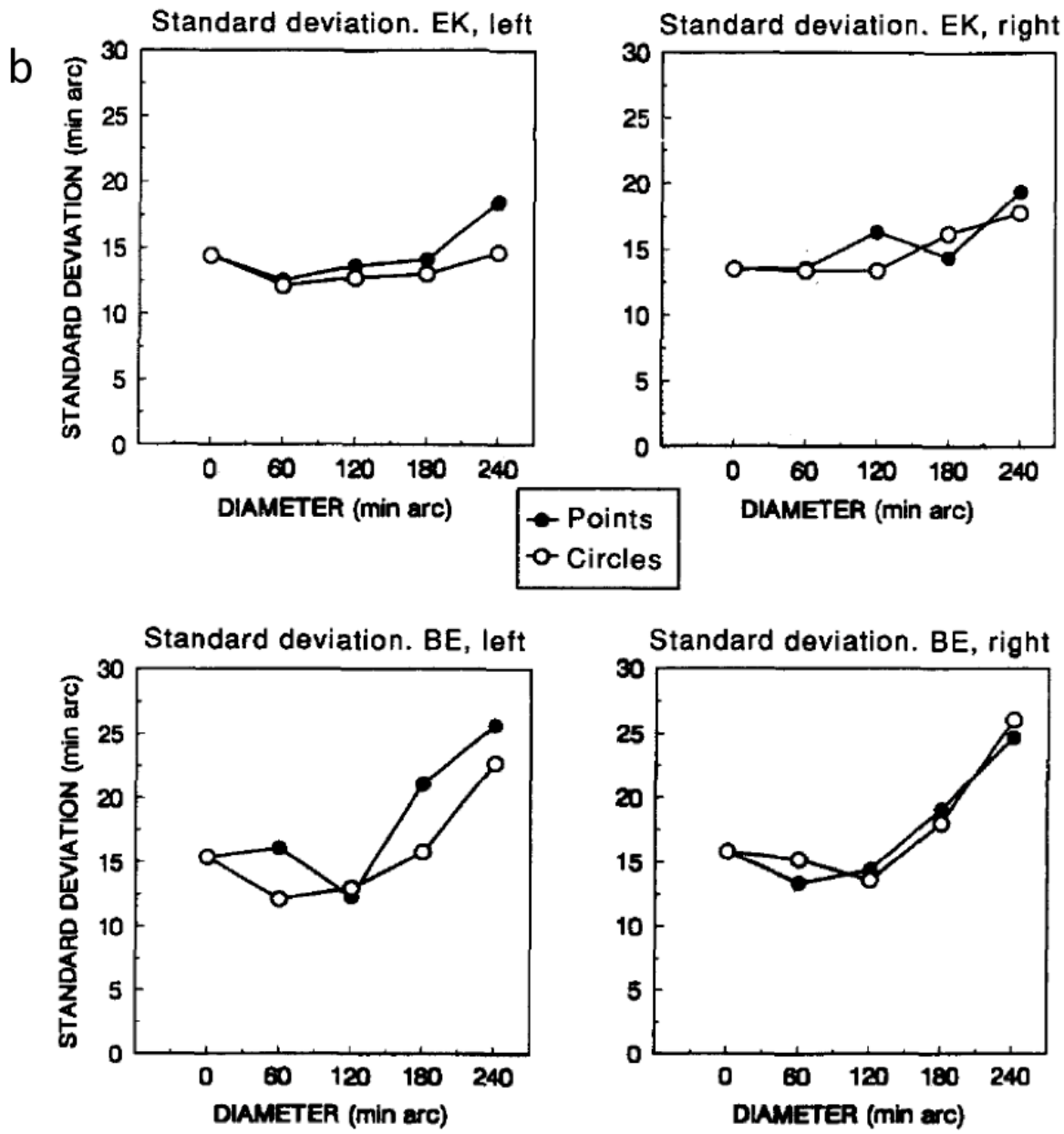


Figure 7.6. (a) Average saccade size in arcminute as a function of target eccentricity for one participant. The top row shows the average responses to diamond (four-point) targets, and the bottom row shows the average responses to circle targets. (b) Mean SDs for the saccade size as a function of target size. The mean SD was calculated by averaging the SD for each of the five target eccentricities. From Kowler and Blaser (1995).

Ploner, Ostendorf, and Dick (2004) also investigated the effects of target size on saccadic eye movements. The stimuli used were red disc-shaped targets (1°, 5°, and 10° diameter) on a grey background that were presented for 50ms. The eccentricities of the targets were between

5° and 15°, with 2.5° increments. Unlike Kowler and Blaser (1995), Ploner and colleagues found that as target size increased, so did saccade latencies. For the 1° target, average saccade latency was ≈218ms, and it increased to ≈232ms and ≈242ms respectively for the 5° and 10° targets. The authors suggested that this difference in results could be due to the different instructions given to the participants. While Kowler and Blaser (1995) specifically instructed participants to ensure accuracy over reaction time, Ploner et al., (2004) did not instruct participants to prioritise either accuracy or reaction time. They further propose that the increased saccade latencies to larger targets were due to the increasingly complex spatial pooling required to accurately land a saccade on the spatially extended target. As for the accuracy of the saccades, it was found that larger targets had saccades that fell further away from centre (i.e. greater undershoot for larger targets), but still to a point within the target. The authors suggest that the ability of the saccadic system to determine saccade endpoints of large targets is akin to that of double-target experiments (e.g. Findlay, 1982 discussed in the following paragraph). For double-target experiments, saccades are computed based on the average of both target positions, whereas for extended targets, information from both ends of a target is used to determine saccade position.

The suggestion by Ploner and colleagues is interesting as Findlay (1982) found that saccadic landing positions are influenced by the presence of a second target. The target "spot" appeared as follows: 1) single target at 2° eccentricity, 2) two targets were at 2° and 3° eccentricities, and 3) two targets appeared simultaneously on both sides of fixation at 2°. For the first two target conditions, the target appeared equally to the left or right of fixation. Participants were required to saccade to the target, and in the two-target condition, fixate on either one of the targets. For both the solitary target and double targets on opposite hemispheres, the saccades were accurately made to the target. In the case of the double targets, saccades were made equally to either the left/right target. For the double targets in the

same hemisphere, the saccade landing position was to an intermediate position between the two targets. Interestingly, the saccade latencies were on average 191ms and 183ms respectively for the single target and double target on the same hemisphere, but were increased to 239ms for the double target on opposite hemispheres. This shows a clear disassociation between the saccadic amplitude and latency measures.

Findlay (1982) took the finding of a saccade landing at an intermediate position between two targets to be indicative of a global effect. This idea was further tested in a follow-up experiment that used square-shaped targets that subtended 0.42° and 0.14°. Participants were required to perform a detection task and a comparison task. For the detection task, participants were required to indicate if a gap was present in the target configuration, and for the comparison task, participants had to indicate if the target configuration was the same as the previous trial. The targets could appear alone at either 5° or 10° eccentricity, or as two targets at 5° and 10° eccentricity. That is, for the double target conditions, it was possible to have both targets the same size (both big, or both small), or different sizes (one big and one small). Coren and Hoenig (1972) found that saccades were directed to the centre of gravity of all stimuli in the vicinity of the target i.e. the presence of an extraneous stimuli influenced the saccade location. Therefore, for double targets with one large and one small target, the saccade landing position should be closer to the larger target. This is precisely what Findlay found. For the same size double targets, the saccade landing position should fall at the midpoint between the two targets. However, Findlay found that saccades did not land on the midpoint, but to a point that is closer to the nearer target. In another experiment, he had targets appear at 2 & 3°, 2 & 4°, 2 & 5°, and 2 & 6° eccentricity. The results were that when the distance between the two targets was small, the saccade lands close to the midpoint. However, when the distance between the two targets was larger, the saccade lands further away from the midpoint (closer to nearer target) of the two targets.

Melcher and Kowler (1999) were also interested in what representation of a target guides a saccade endpoint. They were particularly interested in whether saccades were directed to the centre of gravity (COG), centre of area (COA), or symmetric axis. The COG was defined as the average dot location that made up the boundary of the shape, while the COA was defined as the centre of mass of the shape assuming that the shape is of a uniform density. The symmetric axis is "the locus of points whose minimal distance from the boundary exists to more than one point along the boundary" (p. 2929). The authors used many different shape forms (single dot, circle, ellipse, cardioids, fragmented arcs, curved lines, blobs, and yin). Of all the shapes used, the yin shaped target provided the most interesting results, as the COA, COG, and symmetric axis were located at different points (see Figure 7.7a). Also, the authors found that increasing the size of the circle target from 2° to 3.5° resulted in the mean landing position falling short of the actual intended position (i.e. an undershoot was observed). They claimed that the undershoot was a result of other processes not related to shape processing, and thus adjusted the mean landing position for each participant to take into account this deviation. That is, for each participant, the authors calculated what the undershoot was for a comparably sized circle target, and subtracted this undershoot from the results obtained with other-shaped targets.

The primary finding from the experiments using the yin-shaped target was that saccades were being directed to the COA. To influence the location of the COG, dense clusters of dots were added to the yin-shaped targets (Figure 7.7b). These clusters of dots were superimposed on different areas of the yin shape, and the effect was that the COG (which is a measurement influenced by the density of the dots) was shifted towards the cluster of dots, while the COA (which is a measurement **not** influenced by the density of the dots) remained constant throughout. They found that saccade endpoints were directed to the COA, not COG. The COG was also influenced by changing the boundary spacing of the target (Figure 7.7c). In

this condition, the COG was shifted away from the boundary with dots spaced far apart, while the COA remained the same. Once again the results confirmed that saccades were directed to the COA. The results suggest that saccades were not being directed by the visible elements making up the shape, but instead by an abstract representation of the shape as a whole. The authors claim that this is indicative of saccades being determined subsequent to boundary segmentation. They also argue against the notion of saccades being planned based on the spatial pooling of information, as the cluster of dots did not affect the saccade landing position. It is important to note that the participants were explicitly told to look at the target as a whole, and accuracy was encouraged over reaction time.

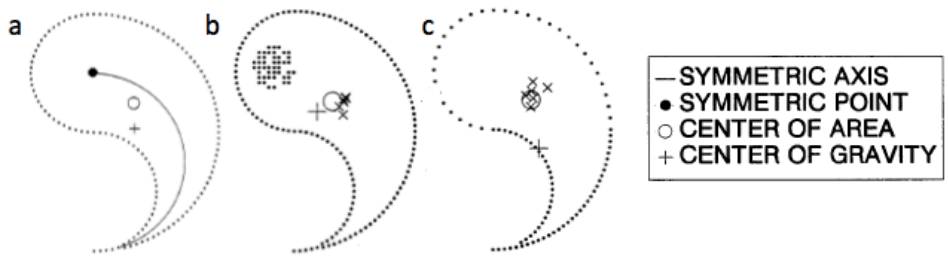


Figure 7.7. (a) A yin shaped target used in the experiment. This shape resulted in different locations for the COA, COG, and symmetric axis and point. (b) Yin shaped target with dot clusters. (c) Yin shaped target with upper boundary dots spaced further apart. The ×s on figure b and c are the adjusted mean landing positions. From Melcher and Kowler (1999).

Table 7.1

Average saccade amplitude and standard deviation in percentage data for various experiments.

Task	Target Type	Target Size	Target Ecc.	% Amplitude (SD)
Kapoula and Robinson, 1986				
Saccade to Target	Red spot	5.6 arcmin	5	97.6 (4.6)
			10	92.9 (2.7)
			15	89.7(5.3)
			20	89.5 (5.1)
Visual Discrimination	E / ∃	0.4° × 0.2°	5	102.4 (5.2)
			10	94.1 (3.1)
			15	87.4 (4.1)
			20	86.1 (4.7)
[a]**Ploner et al., 2004**				
Saccade to Target	Red disc	1°	5° – 15°	91.0 (13.0)
		5°	5° – 15°	88.0 (21.0)
		10°	5° – 15°	84.0 (25.0)
Findlay, 1982				
Saccade to Target	Spot	<1 arcmin	2°	92.9 (15.0)
			2° + 3°	90.6 (12.2)
			-2° + 2°	91.0 (14.4)
Detection	Square	0.14°	5°	117.6 (10.9)
		0.14°	10°	93.7 (7.0)
		0.42°	5°	116.0 (9.2)
		0.42°	10°	98.2 (6.2)
Comparison	Square	0.14°	5°	108.6 (16.5)
		0.14°	10°	96.5 (8.4)
		0.42°	5°	107.3 (12.1)
		0.42°	10°	94.6 (8.6)
[a,b]**Melcher and Kowler, 1999**				
Saccade to Target on Left	Circle	2°	3.8° – 4.2°	98.0 (7.6)
		3.5°	3.8° – 4.2°	94.5 (10.1)
Saccade to Target on Right	Circle	2°	3.8° – 4.2°	105.4 (7.3)
		3.5°	3.8° – 4.2°	110.0 (9.8)

[a] amplitude and standard deviation were estimated from a graph

[b] amplitude and standard deviation were calculated using the average eccentricity of targets

Table 7.2

Average saccade amplitude and standard deviation in percentage data for experiments by Findlay (1982) where two targets were presented simultaneously. The target types were square for all targets.

		Target Eccentricity				
Task	*Target Size*	5° + 10°	2° + 3°	2° + 4°	2° + 5°	2° + 6°
Detection	0.14° + 0.14°	89.2 (10.4)	96.2 (12.0)	84.5 (11.3)	73.8 (14.5)	77.5 (14.1)
	0.42° + 0.42°	89.2 (8.3)	97.3 (9.81)	89.3 (9.11)	77.8 (14.8)	86.5 (17.2)
	0.14° + 0.42°	93.4 (8.1)	96.3 (13.0)	88.6 (10.5)	59.4 (13.9)	73.2 (12.0)
	0.42° + 0.14°	102.9 (11.4)	96.8 (12.9)	92.4 (9.05)	75.4 (11.3)	88.3 (11.6)
Comparison	0.14° + 0.14°	83.5 (9.9)				
	0.42° + 0.42°	85.1 (8.2)				
	0.14° + 0.42°	78.0 (10.1)				
	0.42° + 0.14°	96.9 (14.1)				

Note: amplitude is calculated respective to centre of gravity

7.5 Eye Movements and Texture Stimuli

7.5.1 Role of Eye Movements in Texture Perception

In the earlier introductory chapter, we discussed a number of different studies that investigate texture perception. However, those studies made use of the psychophysical paradigm, where presentation duration of the stimulus is typically limited. As such, the influence of eye movements during the task is reduced, if not completely absent. In this section, we discuss studies that show that eye movements can support texture perception.

He and Kowler (1992) studied how saccades influenced the perception of texture stimuli. Using different figure elements (T, +, and ×) against a constant background element (L), they were able to produce texture patterns that differed in their ease of segregating figure and background regions (Figure 7.8). In a series of experiments, they studied how participants performed in a shape (identify the type of figure elements) and size (identify the number of figure elements) discrimination task. The results show that; 1) performance in both tasks was better when saccades were allowed (scan) compared to when participants had to maintain fixation (stay), 2) saccades (scanning) were especially important for texture patterns that do not segregate easily, whereby for the T vs. L textures, performance reached an asymptote at 1000ms duration for the stay condition, but continuously increased for the scan condition up to the 5000ms duration, 3) saccades do not enhance texture segregation per se, but instead improve the discriminability of the individual target elements that make up the figure and ground, which then aids in boundary segregation, 4) saccades are only helpful if movements bring the fovea to regions of importance (e.g. curvature extrema: minimum and maximum points in the curvature), and 5) increasing stimulus duration (>1-2 seconds) did not improve performance for difficult to segregate patterns (T vs. L) unless saccades are allowed. While the results show that saccades aid performance related to a texture segregation task (size

discrimination) for texture patterns that do not segregate readily, saccades do not actually produce a perceptual border. Performance for the T vs. L pattern under the "scan" condition at stimulus duration of 5 seconds is still poorer compared to performance for the × vs. L pattern under "stay" condition at stimulus duration of 250ms. This shows that a perceptual border is formed relatively easily for the × vs. L pattern, but no amount of eye movements will result in a perceptual border for the T vs. L pattern. Visual inspection of Figure 7.8a vs. c will show this too. At first glance of Figure 7.8c, a border between the figure and background can be perceived instantly, but with Figure 7.8a, no amount of visual inspection will result in a border between figure and background.

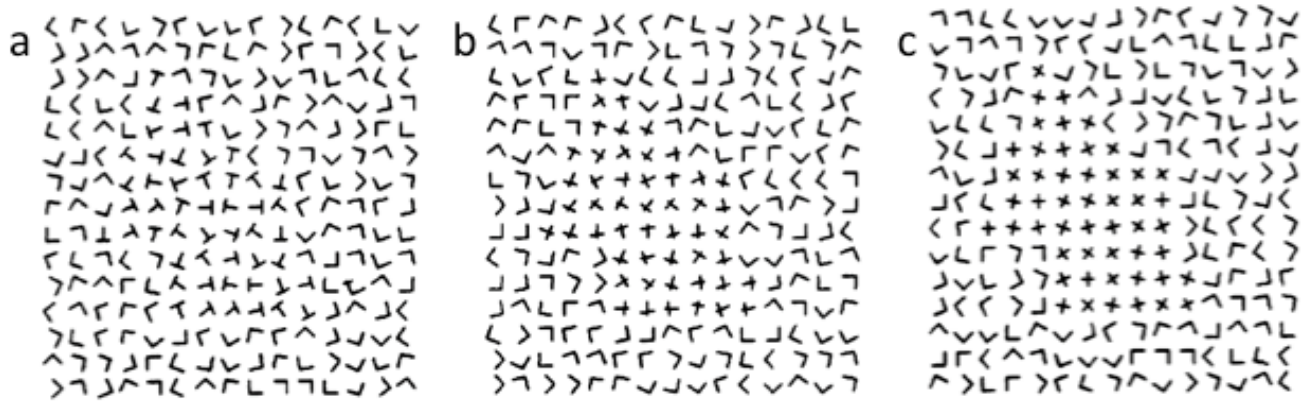


Figure 7.8. Examples of the texture patterns used in the study by He and Kowler. (a) T vs. L, (b) + vs. L, and (c) × vs. L. Visual inspection shows that location the figure region for pattern (a) is not readily noticeable, unlike pattern (c), which has a figure region that easily segregates from the background. The ease of detecting the figure region of pattern (b) lies in between pattern (a) and (c). From He and Kowler (1992).

Van Humbeeck, Schmitt, Hermens, Wagemans, and Ernst (2013) investigated eye movements in contour integration, where the target was embedded in a background of noise elements. Their stimuli consisted of a chain of 7 Gabor elements (the target contour) embedded in a background of randomly oriented Gabor patches. To manipulate the saliency of the target contour, the path angle, which is the angle between successive elements, was

varied at 5°, 10°, 15° or 20°. This is because the smaller the angle between the successive elements, the easier it is for an observer to detect the contour (Field, Hayes, & Hess, 1993 also discussed in Section 2.3). Participants performed two tasks, which were: 1) to indicate if the contour was to the left or right side of the stimulus (LR task), and 2) to indicate if the contour was present or absent (PA task). After a central fixation, the stimulus appeared and remained on the screen until a response was made. The results for both of the tasks show that as path angle increased, accuracy declined while reaction times, number of fixations, and individual fixation durations increased. Besides that, over the course of a trial, the fixation duration increased initially and then started decreasing for both tasks. The overall longer reaction times (owing to the increased fixation count and fixation duration) were due to the additional processing required with decreasing target saliency. Additionally, saccade amplitude measures were not influenced by the path angle, but they were influenced by the tasks. Overall, the LR task produced saccades with larger amplitudes compared to the PA task. However, for both tasks, saccade amplitudes decreased over time, except for the PA task when participants reported the target as absent. This was explained by a coarse-to-fine search strategy, where participants switched from a coarse search (large amplitudes) to a fine search (small amplitudes) when the target was approached.

7.5.2 Saccade Latency to Texture Stimuli

Deubel, Findlay, Jacobs and Brogan (1988) investigated saccadic eye movements to texture (change in orientation) and luminance (change in brightness) targets while varying the presence or absence of distractors and background (Figure 7.9). They found that in the absence of a background structure (Figure 7.9a-b), saccadic latencies were reduced by approximately 30ms (185ms) compared to trials with a background structure. Furthermore, for the trials in which there was a background structure, the modality (luminance or texture)

did not influence the saccadic latencies, which were on average 215ms. They also found that saccade latencies to targets in the presence (Figure 7.9, bottom row) and absence (Figure 7.9, top row) of a distractor were the same. These spatially segregated distractors could have the same (luminance + luminance, orientation + orientation) or different (orientation + luminance) modality as the target.

The results show that while saccadic latency seems to be influenced by the presence and absence of a background, saccadic landing position is not affected by the background structure. Conversely, the presence of a distractor does affect the saccade landing position. Despite participants knowing that there will be a distractor in some trials, their response to the target for trials with a distractor showed that the saccade landing position was determined by averaging the location of both target and distractor to direct eye movement to a point in between the target and distractor. This result demonstrates global processing, where spatial integration occurs during saccadic processing (e.g. Findlay, 1982). Interestingly, spatial integration was found in this study even when the modalities (i.e. orientation / luminance) between target and distractor were different. However, for the orientation-textured modality with background, the orientation-defined distractor had to have an orientation contrast greater than 12° for the global effect to be seen. This implies that the weight of the distractor is taken into account when grouping the distractor and target before averaging. Finally, the spatial integration demonstrated by the global processing implies the use of cells with large receptive fields as saccades are programmed prior to eye movement.

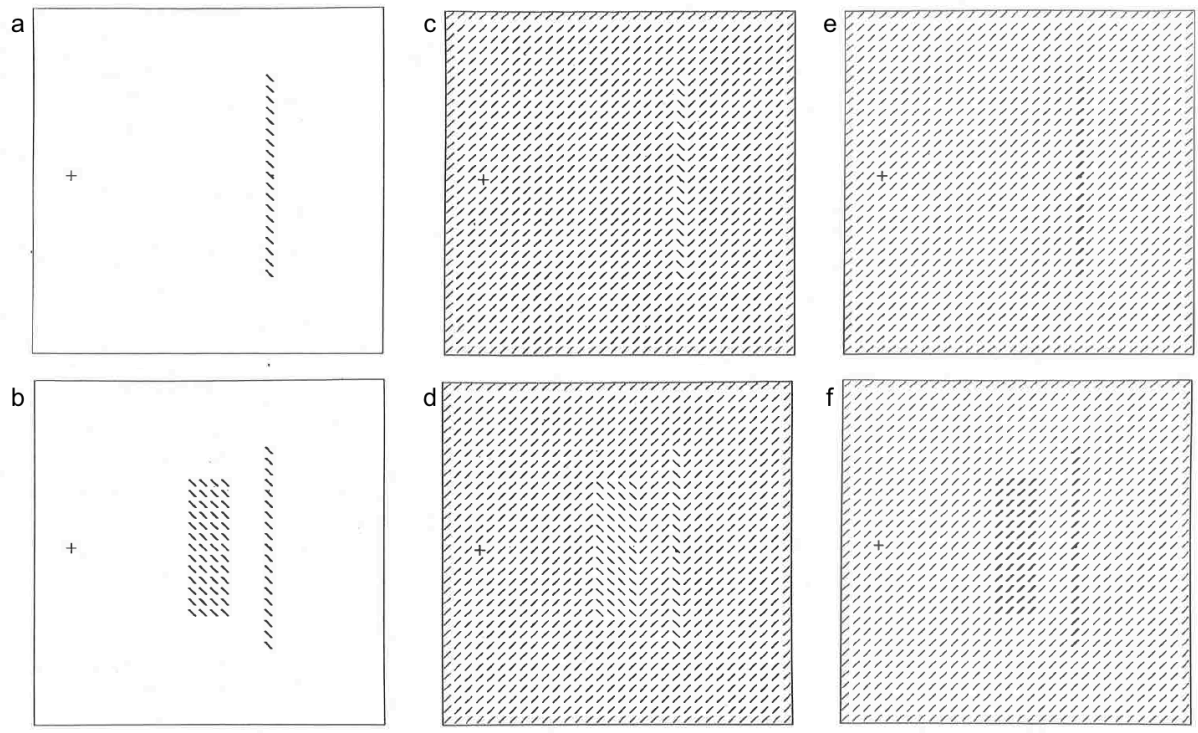


Figure 7.9. Examples of stimuli used in the experiment by Deubel and colleagues. Target was an elongated bar (1X19 line elements), while the distractor was a 4X13 larger region. Top row: Target only stimuli; Bottom row: Target with distractor present. For stimulus with background (c-f), target and distractor comprised line elements that differed in (c-d) orientation, or (e-f) luminance. The distance between the target and distractor (when present) was 2.9°, while the distance between fixation and target could either be 7.4°, 9.5°, or 12.7°. From Deubel et al., (1988).

Deubel and Frank (1991) also studied saccadic latency to texture-defined stimuli with the goal of finding the minimum latency of saccades to targets. To achieve this, they used the gap paradigm, where fixation cross offset occurred 150–250ms before stimulus onset, as saccade latencies were shorter with the gap paradigm (e.g. Fischer & Boch, 1983). When they employed a gap paradigm, in which there was a 200ms gap between fixation cross offset and stimulus onset, saccade latency was decreased by 20-30ms compared to the no gap paradigm where saccade latency was 185ms. Furthermore, they also found that saccades to targets alone (Figure 7.10a) took approximately 200ms (180ms with gap paradigm), and this

increased by 30-40ms (240ms, and 220 with gap paradigm) in the presence of a background structure (Figure 7.10b). These results are in line with the findings of Deubel et al., (1988), who found that saccade latencies to targets with a background structure are longer compared to solitary targets with no background. Deubel and Frank conducted a further experiment using a double-step paradigm. As per the aforementioned experiment, participants had to make a saccade to a texture-defined target, which could appear in the presence and absence of a background structure. However, with this paradigm, the initial target will appear (Figure 7.10c), and after a variable delay in time, the initial target will be replaced with a new target at a different location (Figure 7.10d). For the most part, the target was shifted before a saccade was made. The results showed that participants made saccades to the new target if the time between target shift and saccade onset was at least 90ms for the no background condition and 160ms for the background condition. This implied that when a target is embedded in a background, the increase in saccade latency to the target is due to the additional time needed by the visual system to first segregate figure from ground, and this process of localizing texture borders takes at least 160ms. On the other hand, the shorter latencies associated with the gap paradigm is most likely due to a temporal warning that is received by the offset of the fixation (Rayner, 1998).

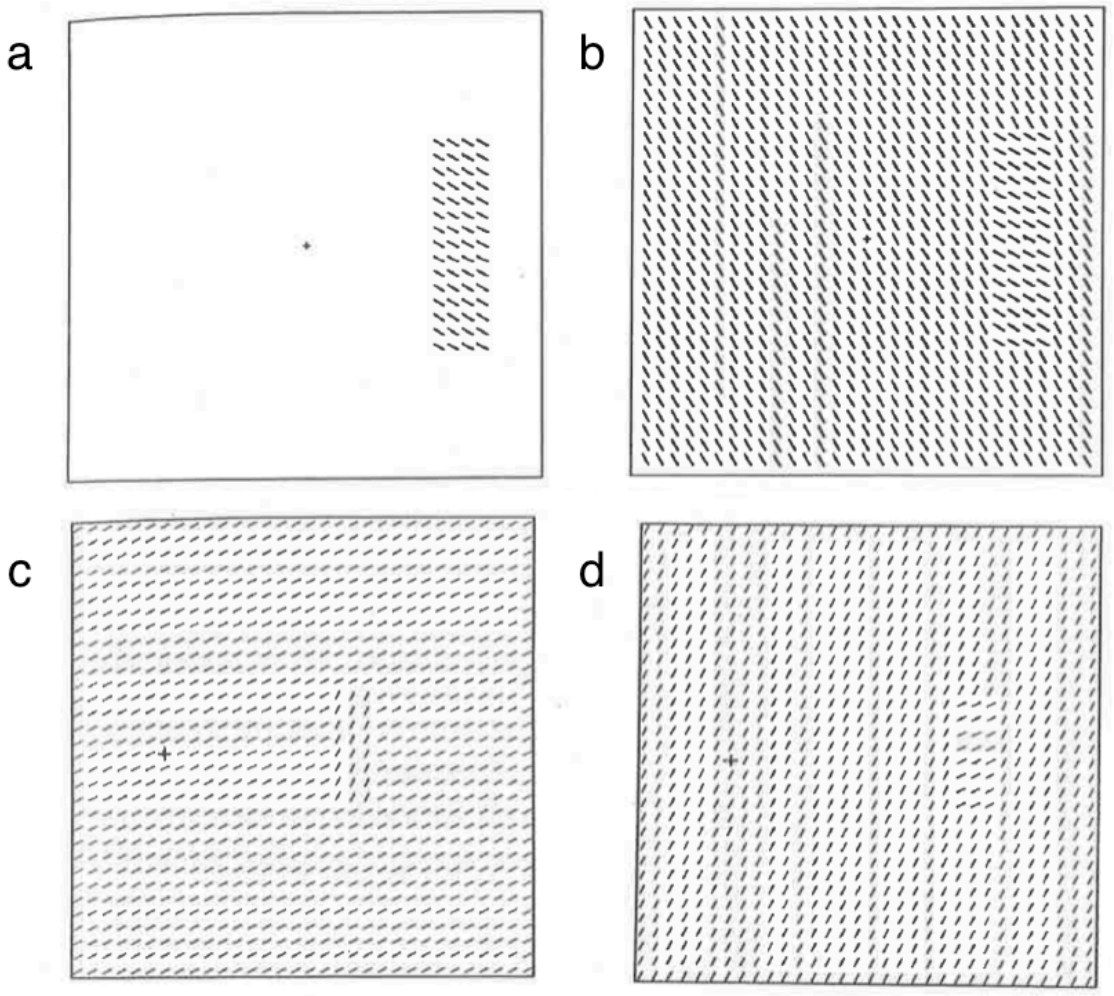


Figure 7.10. a) Target with no background condition. b) Target defined by orientation contrast from the background structure. Bottom images: double-step paradigm with background structure. c) The initial target displayed, and d) the new target display after a variable delay. The targets could appear on either side of the fixation, at eccentricities of 5, 5.5 and 6 degrees. From Deubel and Frank (1991).

Additionally, in the interest of finding the absolute minimum latency of a saccade, Deubel and Frank (1991) conducted pilot studies using the gap paradigm with stimuli made up of a target defined by a change in orientation from the background. They found that an orientation difference between 30-40° produced the shortest saccade latencies, with smaller and larger changes in orientation producing longer latencies (Figure 7.11). Considering how most research (e.g. Lamme et al., 1999; Roelfsema, Lamme, Spekreijse, & Bosch, 2002; Supèr,

2006) that studies the neural mechanisms of figure-ground segregation use stimuli with orientation change of 90°, this is certainly an interesting point to note. This is because Deubel and Frank's finding suggests that the underlying neural mechanisms for texture segregation are optimal (as indicated by the shortest latencies) when orientation contrast is between 30°-40°, which could have an impact on studies that use textured-stimuli with 90° orientation contrast. Furthermore, the experiments described above had participants fixate at a cross in the middle, with targets appearing on either the left or right side. Using this paradigm, they were able to get cut-off (minimum latency) durations, in which anything lower than the minimum latency produced anticipatory saccades. However, when they piloted the study using an experimental paradigm where the target position was not varied (i.e. target always appeared to the right of the fixation), it was impossible to distinguish anticipatory saccades from regular saccades, because the directions of the saccades were always correct.

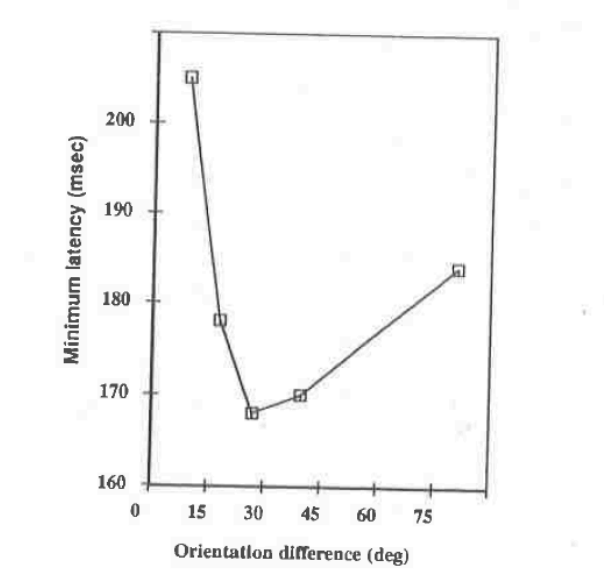


Figure 7.11. Saccade latencies to texture-defined (with background) targets, with the gap paradigm (gap =200ms). Saccade latencies are plotted as a function of orientation contrast. From Deubel and Frank (1991).

Nothdurft and Parlitz (1993) also studied the occurrence of express saccades (i.e. very fast, <100ms) by recording the distribution of reaction times for saccadic eye movements to targets. These targets could be defined by a change in orientation, motion, or luminance (Figure 7.12a-c). They found that orientation defined targets did not produce express saccades, as the distributions of the saccadic reaction times (SRTs) were unimodal (one peak), with peaks well above 100ms (140-220ms). Distribution of SRTs for the motion defined targets were similar to the orientation defined targets, but not to the luminance defined targets that produced a bimodal distribution (two peaks) with the latency of the first peak occurring before 100ms (70-105ms) and the latency of the second peak occurring after 100ms (120-170ms). The results indicate that only targets defined by luminance, not orientation or motion, produce express saccades.

However, it is possible that orientation targets did not produce express saccades because the signal strength was too weak. To test this possibility, Nothdurft and Parlitz conducted a matching experiment in which participants were presented with two targets simultaneously (Figure 7.12d-e), and had to saccade to the most salient target. For the orientation-luminance match, the orientation of the single line element was held constant at 90°, while the luminance contrast was systematically varied. From that data, they were able to get a point in which luminance contrast had an equal saliency to the orientation target. With that particular level of luminance contrast, the distributions of participants SRTs were unimodal with peak latencies between 170-220ms, which is similar to performances with orientation defined targets. The results show that express saccades are only elicited when target contrast is high enough to invoke them, and orientation cues at maximum contrast of 90° are not salient enough to invoke express saccades. Furthermore, SRTs to orientation targets were also slower compared to the regular (not express) saccades for luminance contrast. This implies that luminance contrast targets are more salient, which in turn elicits faster saccades.

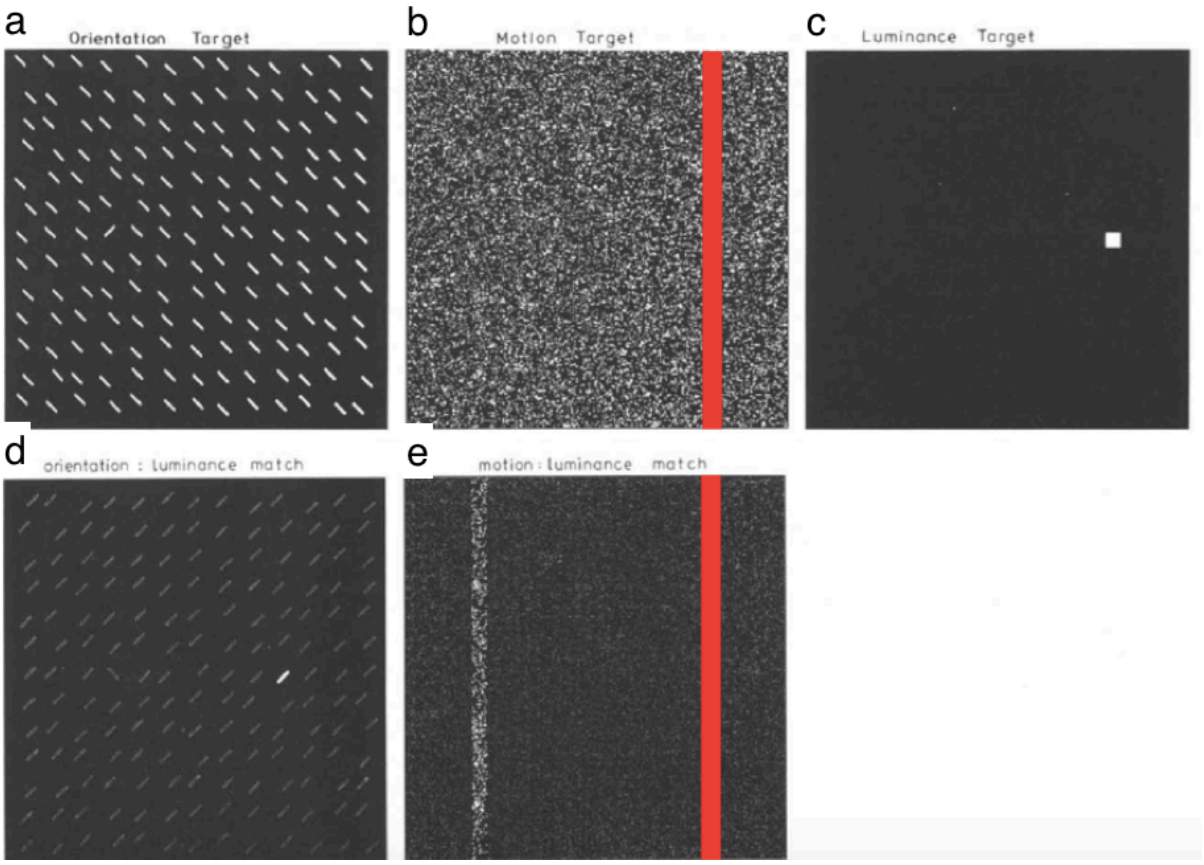


Figure 7.12. Examples of the texture patterns used in the study by Nothdurft and Parlitz. The targets could appear on either the left or right side of the fixation (centre of stimulus). (a) Orientation defined stimuli consisted of a single line element that is perpendicular (90°) to the rest of the lines. (b) Motion-defined stimuli consisted of small dots moving either up/downwards, and a target (red bar as a visual representation, with a visual angle of 0.8° wide) of dots moving in the opposite direction. (c) Luminance-defined stimuli consisted of a bright square against a dark background. For the matched experiments, the stimulus will display two targets simultaneously on either side of the fixation point. In these tasks, orientation (d) and motion (e) defined targets are included with luminance defined targets. From Nothdurft and Parlitz (1993).

7.5.3 Strength of Neural Response

In the previous section, we discussed studies that investigate the time it takes to make a saccade to a texture stimulus. In this section, we discuss studies that investigate the neural activity in the visual cortex that produces the oculo-motor response to the texture target. This

neural activity is shown to be dependent on the properties of the stimulus. The reader is also directed to Section 2.2 that discussed the neurophysiology of texture perception.

Supèr, Spekreijse, and Lamme (2001) analysed neural responses in V1 of rhesus macaques while they were engaged in a figure detection task. The monkeys were trained to make a saccade to a figure, which was made of a square region with lines oriented orthogonal to the background lines (90° orientation contrast). On some trials, a uniform texture was presented, and the monkeys had to maintain their fixation. The monkeys were able to correctly detect the presence of the figure for 78% of the figure-present trials ('seen trials'), and of the 22% trials when they did not make a saccade to the figure, they maintained fixation for 8.1% of those trials ('not seen trials'). The experimenters were primarily interested in the difference of neural responses for the 'seen' and 'not seen' trials. They found that both 'seen' and 'not seen' trials exhibited neuronal responses in the V1, but the 'seen' trials exhibited contextual modulation while the 'not seen' trials did not. This contextual modulation refers to the difference in neuronal responses between the figure and ground region of a texture, where signal strength of the figure region is greater (e.g. Lamme et al., 1999). The results indicate that during 'not seen' trials, early V1 activity (before 90ms) is still evoked, but signal strength of the figure region is not greater than the ground region 100-240ms after stimuli onset.

Supèr, Spekreijse, and Lamme (2001) also recorded neural responses while manipulating the saliency of the stimulus by decreasing the line length (see Figure 7.13a). With shorter lines, the detectability of the figure also decreased (e.g. Nothdurft, 1985b), and the authors were able to analyse the neural activity during 'seen' and 'not seen' trials. For all line lengths, early V1 activity was exhibited, but with decreasing line length, contextual modulation had a weaker effect on 'seen' and 'not seen' trials. That is, with short line lengths, contextual

modulation for ‘seen’ and ‘not seen’ trials were similar to each other. This indicated that contextual modulation by itself could not explain why some trials are ‘seen’ and ‘not seen’. The authors proposed a model (Figure 7.13b) in which the strength of internal representation (third column) produced a ‘detect’ and ‘non-detect’ mode. The internal representation was theorized to be due to certain inherent factors, such as the state of the brain during stimulus presentation, that results in an internal representation to be generated or not. Therefore, regardless of saliency of stimulus, a certain percentage of trials will not be detected due to internal representation, and trials that were not detected will always produce ‘not seen’ responses. However, for stimuli with high saliency, the decision processing stage (fourth column) will always yield a ‘seen’ response as the signal strength is always stronger than the threshold of the response criterion. For the low saliency stimuli on the other hand, the strength of the signal is close to the threshold of the response criterion, which results in some trials being passed as ‘seen’ responses, while others are ‘not seen’ responses. This model accounts for why contextual modulation has a strong correlation with ‘seen’ and ‘not seen’ trials for highly salient stimuli only.

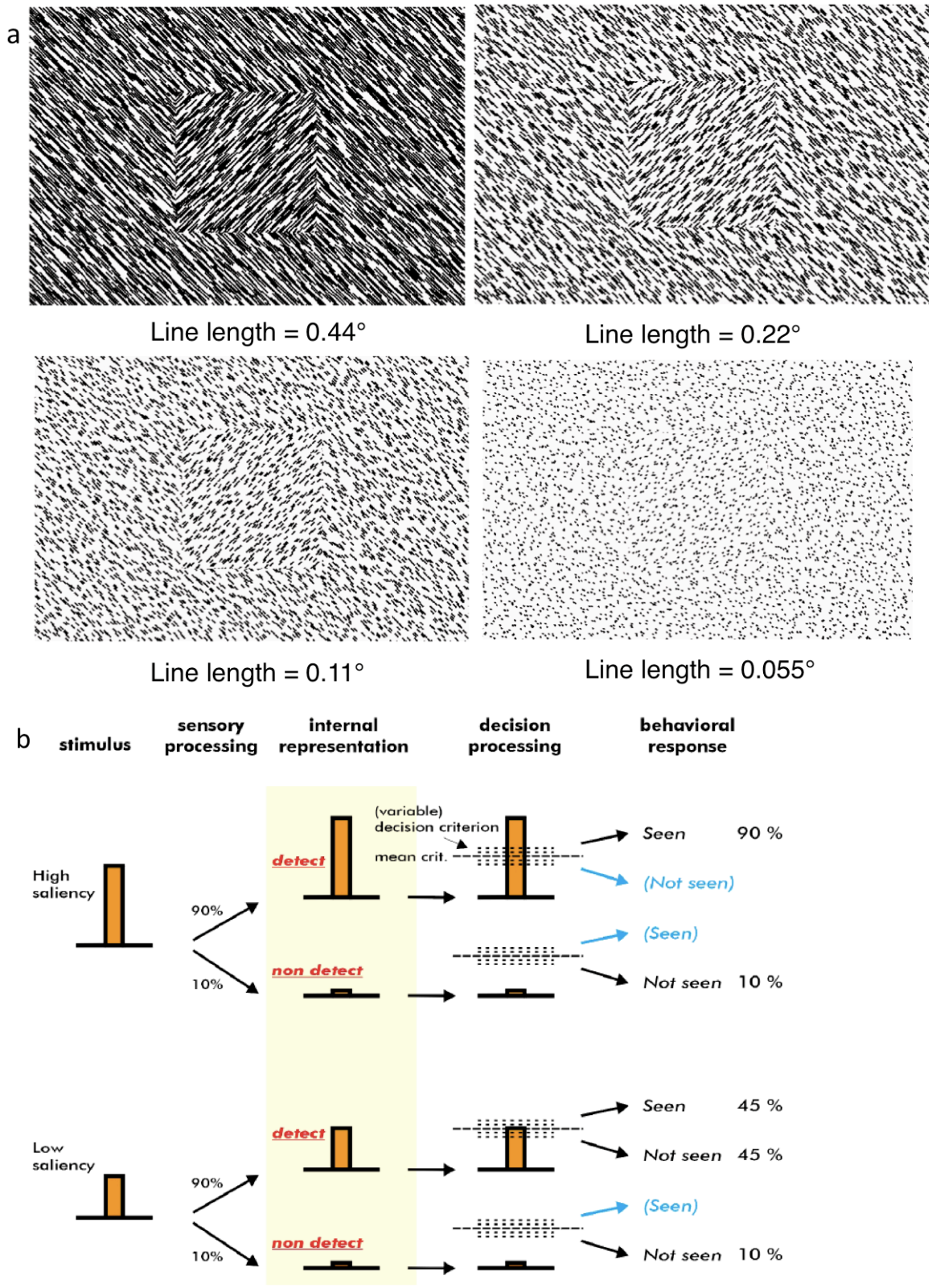


Figure 7.13. a) Different line lengths were used to manipulate the saliency of the figure, b) Schematic diagram of the model proposed by Supèr and colleagues. From Supèr, Spekreijse, and Lamme (2001).

With regards to the internal representations mentioned above, a follow-up study on the internal state of the monkey while performing a detection task was conducted (Supèr, van der Togt, Spekreijse, & Lamme, 2003). The experimental setup was the same as above, but with the inclusion of a pre-stimulus screen made up of randomly oriented lines being presented to the subjects 300ms before stimulus onset. This was primarily because neural activity in the V1 was also being recorded 300ms before stimulus onset to gauge the influence of the internal state of the V1 on detection performance. They found that functional connectivity (as measured by the correlation of signal strength between multiple electrodes) between neurons for 'seen' trials were greater than 'not seen' trials. Moreover, neural activity 100ms before stimulus onset (-100 to 0ms) was stronger for 'seen' trials compared to 'not seen' trials. However, unlike the previous study which showed a difference in figure and ground signal strength (contextual modulation) during late activity (>100ms) for 'seen' trials (Supèr, Spekreijse, & Lamme, 2001), there was no difference in signal strength of pre-stimulus activity (-100 to 0ms) to figure and ground region. Despite this difference in contextual modulation between the pre-stimulus and post-stimulus (>100ms) activity, the overall stimulus-evoked responses were highly correlated with each other. In contrast, the correlation between early and late activity (i.e. 0 to 100ms vs. >100ms), as well as pre-stimulus and early activity (i.e. -100 to 0ms vs. 0 to 100ms) were not strongly correlated. Overall, the results suggest that for figures to be reported, neural activity in the V1 100ms before stimulus onset has to be both synchronous and strong.

Supèr, Spekreijse, and Lamme (2003) extended their earlier findings to study the effects of signal strength on saccade latency. Using the same experimental setup as described above (Supèr, Spekreijse, & Lamme, 2001), they grouped together saccadic reaction times into early ($\bar{x}$ = 260ms) and late ($\bar{x}$ = 320ms) latencies, and compared the strength of contextual

modulation (stronger neuronal response to figure then ground) for both groups. They found that early V1 activity was the same for both groups, but late activity (>100ms) showed an increase in contextual modulation for the early latency group compared to the late latency group. Further analysis showed that the onset of contextual modulation for both groups was approximately 100ms. This indicates that the strength, not onset, of contextual modulation affects saccade latency. And considering how contextual modulation has been linked with horizontal and feedback connection (Lamme, Supèr, & Spekreijse, 1998), impairment or delay in these recurrent interactions could result in the variability seen in saccade latencies.

However, the studies discussed thus far recorded neural strength either before or during stimulus presentation. As a result, the neural responses recorded could possibly also reflect motor-related processing as the neural activity to the stimulus coincided with preparation of the saccade. Therefore, with the studies discussed previously, it is unclear whether V1 activity was a reflection of sensory or motor processes. To address this, Supèr and Lamme (2007) carried out a similar study as above, but with a delayed detection task, in which the monkeys made eye movements to the figure 1000ms after stimulus onset. The experiment also employed two different conditions: visually-guided (figure remains on screen) and memory-guided (figure changes to uniform stimulus) trials. They were interested in studying how V1 activity related to saccade latency and direction of saccade.

They found that memory-guided trials displayed a negative correlation between signal strength and saccade latency (i.e. increased neural response results in faster saccade latencies), while visually-guided trials did not. Moreover, the negative correlation is only found in response to figure locations only, not background, which implies a certain degree of spatial specificity. The negative correlation between signal strength and saccade latency becomes significant from ~89ms onwards, which is comparable to the time contextual

modulation begins (e.g. Lamme, 1995; Lamme et al., 1999). Finally, both visually and memory-guided trials show an increase in contextual modulation just before saccade onset. The outcome of this study is that V1 activity can guide a saccade by planning when (saccade latency) and where (direction of saccade) a saccade is made. This process involves a neural response by the sensory system after stimulus onset, which then produces a build-up in the motor system prior to saccade onset. When a strong figure-ground signal is produced, the build-up in the motor system will be rapid, resulting in the threshold being reached earlier, which in turn produces faster saccades.

The key outcomes of these studies are that: 1) the neural strength of a population of neurons, as well as the connection between them, is necessary to perceive a figure-ground texture; 2) Increased neuronal activity to the figure region compared to the ground region (contextual modulation) is important to segregate a texture stimulus; 3) The stronger the contextual modulation, the faster the saccades; 4) The importance of contextual modulation is reduced for stimuli with low saliency; 5) Faster saccades are elicited when the strength of the neural response to the figure location is increased. In sum, these studies imply that figure-ground segregation has to occur before a saccade can be planned. In fact, the stronger the neural activity (to indicate figure-ground segregation), the faster a saccade can be evoked. Thus, if a stimulus is not very salient, the neuronal activity will be weak, and saccade latencies will be longer.

7.6 Motivation for Eye Tracking Experiments

Most research on saccade control has centred on simple stimuli with single/isolated targets that appear in an impoverished visual environment (i.e. no background structure), or complex stimuli of natural scenes i.e. high-level stimuli. As far as we know, there are only a limited number of studies that investigate saccade control to low-level stimuli where targets are

embedded in a textured background. However, the studies that do investigate eye movements to figure-ground texture stimuli have used target figures that were spatially small. Therefore, it is currently not known how a saccade destination will be programmed for spatially extended targets that are embedded in a background structure. This was one of the aims of the experiments presented in the following chapters.

Additionally, the configuration of the target figures in previous studies was always very similar – either square or rectangular shaped targets that have an abrupt texture difference at the edge between the figure and background. As we discussed in Chapter 2, a texture edge is not necessarily crucial for figure-ground segregation (e.g. Kingdom & Keeble, 1996; Ben-Shahar & Zucker, 2004). Yet, studies that investigate eye movements to texture stimuli have always used stimuli where the texture property at the edge changes abruptly. Thus, it does not inform about how saccades are programmed for texture stimuli that do not have an abrupt texture edge. This was another aim of the experiments conducted.

Along the same line as our choice to use texture stimuli with an abrupt and smooth texture variation, the experiments that will be presented here are meant to be an **extension** of the psychophysical work conducted. That is, we are still interested in investigating if texture edges are crucial for figure-ground segregation (the overarching objective of all the psychophysics experiments). However, we are now using a different methodology that will hopefully gives us **more** information as we are not limiting the contribution of eye movements in texture segregation.

Moreover, the eye movement measures reported in previous studies of texture segregation were saccade amplitude and saccade latency, which were respectively the landing position of the first saccade, and the time it took to initiate the first saccade. However, other eye movement measures could also provide insight into information processing during texture

segregation. Thus, in the experiments presented in the following chapters, we use other eye tracking measures such as fixation count and summed fixation duration.

We present two experiments in Chapter 8, one of which was a pilot study, in which we recorded participants' eye movements while they searched for a figure embedded in a background. Unlike the studies discussed in Section 7.5, we use target figures that were large i.e. spatially extended, and also had different figure-ground orientation profile configurations. These profiles were the same as used in the psychophysics studies presented in Chapters 4 – 6. For the pilot study, participants were only required to look for the figure, and not perform any other tasks, whereas for the follow-up experiment, they were required to localise the edge and centre of the texture figure. The purpose of these experiments was to investigate the eye movements associated with a texture segregation task, especially with conditions where the texture figure was easy and difficult to locate.

However, the design of the experiments in Chapter 8 was not conducive to address all of our research questions, specifically in regards to saccade destination. This was addressed in Chapter 9, where we presented experiments with a modified experimental paradigm that was suited to answering our research question. To do this, we only used texture stimuli where the figure was easy to detect, and could appear at randomly selected eccentricities to the right or left of a central fixation. This setup was designed to promote saccades made directly to the target.

For both Chapters 8 and 9, eye movements were studied in response to stimuli with different profile configurations. However, this did not directly inform us about how **salient** the different profiles were. Thus, in Chapter 10 we presented a single experiment where we wished to match the effectiveness of the various profiles in attracting an eye movement. Instead of just presenting a single figure embedded in a background, participants were

simultaneously shown two figure-ground texture stimuli on either side of a central fixation point, and they had to make a saccade to the texture figure that was most noticeable to them.

Chapter 8

Investigation of Eye Movements to High and Low Orientation Contrast Texture Stimuli

From the literature discussed in Chapter 7, particularly in reference to eye movement studies using texture stimuli, it is clear that very few studies investigate eye movements to texture stimuli. The few studies that do investigate this tend to focus more on saccade latency (e.g. Deubel & Frank, 1991; Deubel et al., 1988; Nothdurft & Parlitz, 1993) or the strength of neural responses in the visual cortex (e.g. Supèr, Spekreijse, & Lamme 2001; 2003). Hence, there is a gap in the literature pertaining to where in a texture stimuli participants would look.

The aim of the studies presented in this chapter was to examine the involvement of eye movements in the process of texture segregation. We were particularly interested in investigating how eye movement patterns differ between textures with figures that pop-out easily and those that do not. We were also interested in how different orientation profiles – Block, Blur, and Cornsweet – would influence patterns of eye movement.

In the pilot study that begins this chapter, participants were shown texture stimuli for a fixed duration, and were told to locate the figure within. Specifically, they were told to look for the figure within the time given, without having to make any other response. Then, in Experiment 7, participants were once again shown texture stimuli, but they were asked to perform one of two tasks – to either locate the centre or the edge of the figure. For both the pilot study and Experiment 7, participants' eye movements were tracked while they looked at the texture stimuli. The eye movement measures from these studies could allow us to draw conclusions about information processing during texture segregation.

8.1 Pilot Eye Tracking Study with Texture Stimuli

As discussed in Section 7.5, studies that investigate eye movements to textured targets typically use stimuli where the entire figure region consists of a uniform orientation that differs in orientation from the ground region i.e. with an abrupt texture edge. However, psychophysical studies (e.g. Kingdom & Keeble, 1996; Ben-Shahar & Zucker, 2004) have shown that texture segregation is possible even when the texture property varies in a gradual manner, i.e. **without** an abrupt texture edge. As far as we are aware, no studies have looked at eye movements involving these stimuli with smooth texture variations.

Thus, one aim of this pilot study was to determine if eye movements to texture stimuli are affected by the different orientation profiles. Stimuli with different texture configurations were used, where orientation change between figure and ground could form three different profiles (see Section 3.3 for more details). Briefly, the Block profile exemplifies stimuli with abrupt texture variations, while the Blur profile exemplifies stimuli with smooth texture variations. There was also the Cornsweet profile that had an abrupt texture change at the edge, which tapers off further away from the edge i.e. the orientation of the elements was the same within the centre of the figure and the ground region.

In addition to the different orientation profiles, the amount of orientation contrast between figure and ground was manipulated, which influenced the saliency of the figure. When orientation contrast is high, the figure patch is more salient and segregates easily. On the other hand, when orientation contrast is low, segregation is not as easy, and the figure does not pop-out from the background. This manipulation was to address the second aim of the study, which was to investigate if eye movements are affected by how easily the figure segregates from ground.

During analysis, Areas of Interest (AOIs) were created to examine where in a texture profile participants' eye movements would predominantly be directed at (see Figure 3.10 in Chapter 3 for AOIs created). An eye movement measure we were particularly interested in was the region of first fixation (RoFF), which informs where in a texture stimulus participants initially look. Moreover, we used measures of total fixation count as an indication of saliency or importance, and also as a measure of difficulty (Holmqvist et al., 2011). The summed fixation duration measure was used as an indication of high levels of interest to a region (e.g. Janik et al., 1978), as well as greater cognitive exertion during information processing (Rayner, 1998).

For both the total fixation count and summed fixation duration, we hypothesised that participants would fixate on regions of a texture figure that produces the most information regarding orientation contrast between figure and ground. As such we might expect fixations to be directed to the centre region of the figure for the Blur profile, and edge region of the figure for the Cornsweet profile. For the Block profile that has information regarding orientation contrast at both figure centre and edge, we might expect that fixations would be equally distributed between both regions.

We also hypothesise that stimuli with low orientation contrast will have a greater number of fixations than high orientation contrast stimuli owing to the reduced detectability of the figure. Likewise, as fixation count is indicative of difficulty in processing a stimulus, we hypothesise that the number of fixation will be influenced by the different orientation profiles. From the psychophysics experiment conducted, performance as measured by orientation contrast threshold consistently found that the Block profile segregated most easily (Block > Cornsweet > Blur). Hence, we expect a similar pattern here, where the number of fixations will be highest for the Blur profile, followed by the Cornsweet, then Block profile.

This study was largely exploratory in nature. In fact, only after conducting this study and Experiment 7 (also presented in this chapter) were we able to conclude that a different methodological approach would have been more informative in addressing the research question in regards to the location of the first fixation. This will be addressed in greater detail in Experiment 7.

8.1.1 Methods

8.1.1.1 Participants

Twenty-four participants (M_{age}= 20.8, SD_{age}= 1.4, 5 females) were recruited from the University of Nottingham Malaysia. All participants had normal, or corrected-to-normal vision.

8.1.1.2 Apparatus & Display

A Tobii T60 Eye-Tracker running Tobii Studio 3.2 was used to present the stimuli and record participants' eye movements. A detailed specification of the setup is presented in Section 3.6.

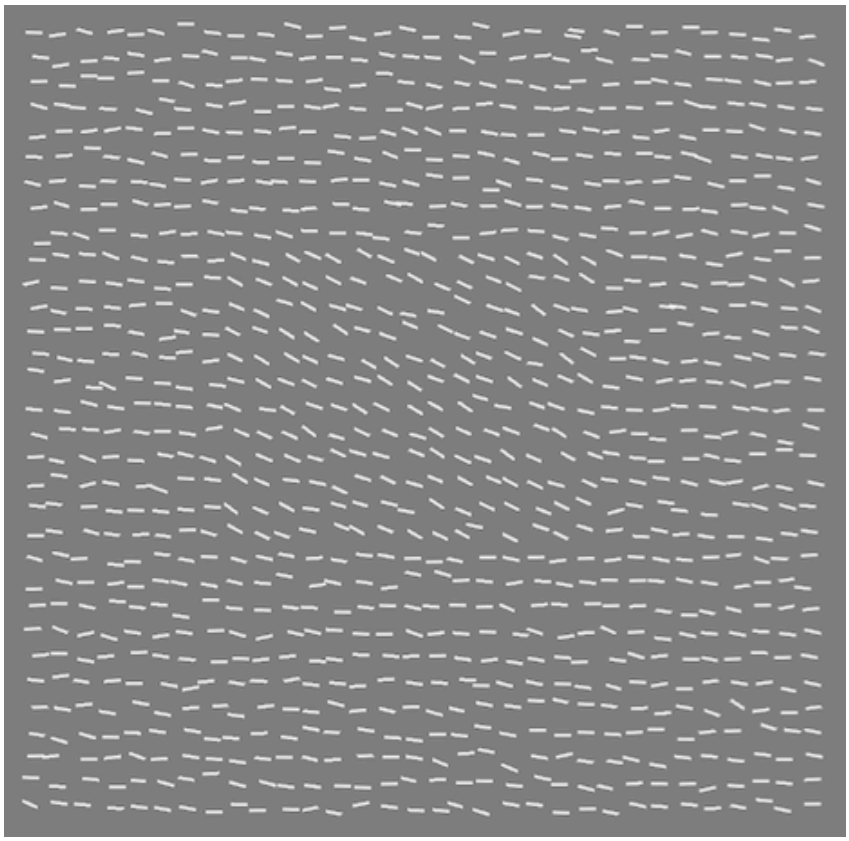
Participants sat 65cm away from the monitor, which produced a resolution of 1.4 arcminute per pixel. The display screen was 1024×768 pixels, which spanned 23.9°×17.9° of visual angle. The 32×32 texture grid had a visual angle of 10.9°, while the 15×12 figure patch was 5.3°×4.2°. The distance between the fixation cross and the middle of texture grid was 12.0°.

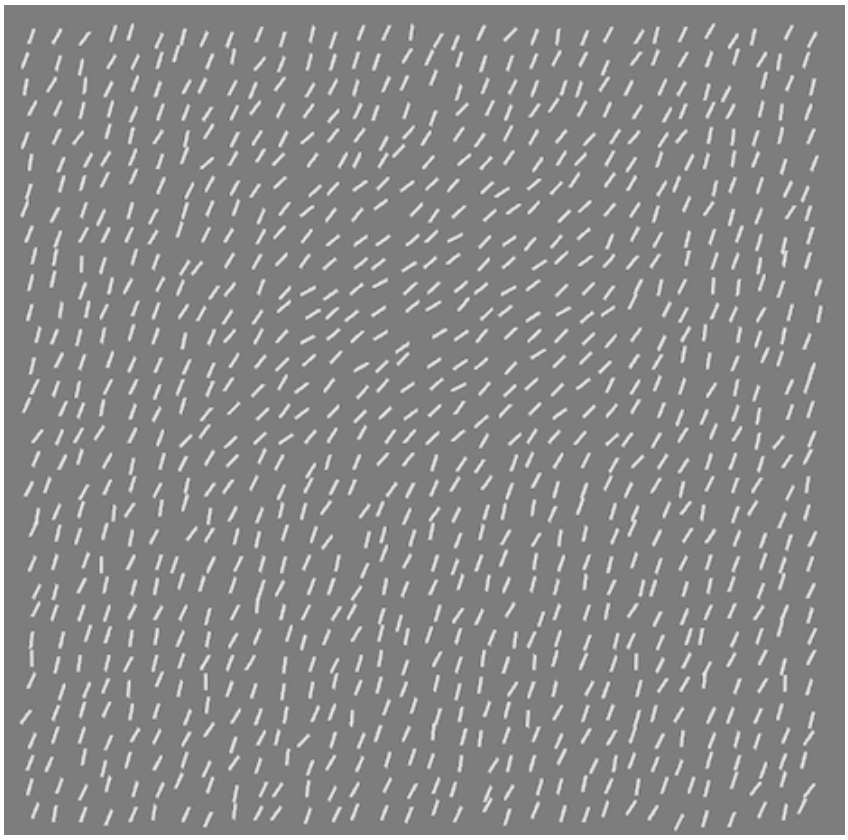
8.1.1.3 Stimuli

The stimuli were pre-generated using PsychoPy 1.83.04 (Pierce, 2007) on an Apple Mac Mini. They were created in the same manner as described in Section 3.3.

The stimuli either had low or high orientation contrast. For the high orientation contrast stimuli, the target figure would be easy to detect whereas for the low orientation contrast stimuli, the target figure would still be detectable, but it would not pop-out and be easy to locate. The orientation contrast was set to 90° for high contrast stimuli, as that is the highest possible degree of orientation contrast that is physically meaningful. For the low contrast stimuli, the values were set around threshold, which is the level at which the presence of a difference between figure and ground is just detectable (i.e. at 82% correct in a 2AFC detection task, with stimulus displayed for 216ms). They were 19° for the Block profile, 33° for the Blur profile, and 29° for the Cornsweet profile. The orientation contrast values were chosen based on the threshold of the naïve observer who participated in Experiment 1.

There were six conditions (orientation contrast – high and low; orientation profiles – Block, Blur, and Cornsweet), with five examples of stimuli per condition, totalling 30 stimuli for the entire experiment. Figure 8.1 shows examples of stimuli for the Block, Blur, and Cornsweet profiles at low orientation contrast. See Figure 3.2 in Chapter 3 for examples of stimuli with 90° orientation contrast. Three areas of interest (AOIs) were created for all stimuli (see details in Section 3.6.2) – specifically the centre, edge, and ground.





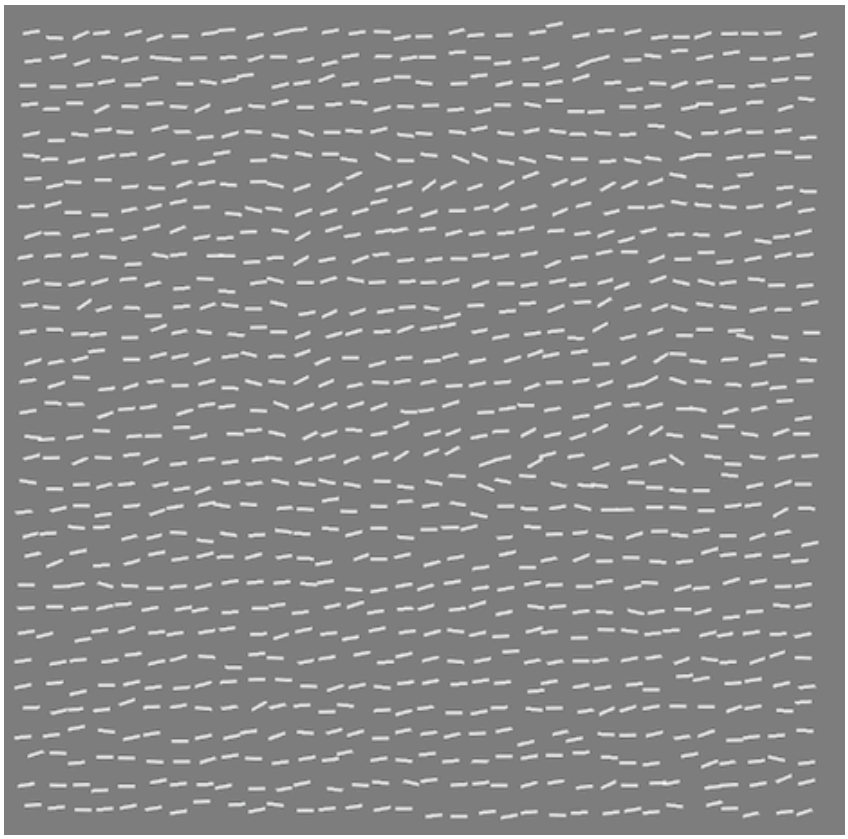


Figure 8.1. In order from the top, example of the Block (19°), Blur (33°), and Cornsweet (29°) stimuli. Orientation contrast is shown in parentheses.

8.1.1.4 Procedure

Prior to starting the experimental session, the experimenter showed the participants six examples of stimuli (one for each condition). These examples were different to those that the participants were actually shown in the experimental session i.e. not repeated. The participants were shown these examples because they were not familiar with texture stimuli, and the verbal description of 'target figure' might be unclear to them. The participants were shown these examples of stimuli on the eye tracker monitor, and the target figure was explicitly pointed out to them. Participants had an unlimited amount of time to familiarise themselves with the stimuli, but were told that in the experimental session, the texture stimuli would only remain on-screen for 5000ms.

Our global aim was to investigate the pattern of eye movements elicited in response to texture stimuli. Because this study was exploratory in nature, we were interested in what would occur when participants were told to look for a figure i.e. a naturalistic response without a prompted outcome. Specifically, participants were instructed: "*once the textured image appears on-screen, all you need to do is look for the small figure inside the bigger square. You will have five seconds to do this before the image disappears"* i.e. they were not told to look at any specific point within the target figure, nor were they required to perform any other response other then locating the figure visually. As a result of these instructions, there was no quantifiable measure of whether or not the participants were actually able to locate the target figure. Furthermore, no further instructions were given about where participants should look once the figure was located i.e. they were not told to maintain fixation at the figure region.

The sequence of a trial began with a fixation cross presented at the left side of the screen for 1000ms. Immediately after fixation offset, the texture stimuli was presented for 5000ms. All 30 stimuli were presented only once per participant in randomized order, and every participant completed the experiment in one session.

8.1.2 Results

Eye movement measures analysed in this study were the region of first fixation (RoFF), fixation count, and summed fixation duration. However, as mentioned above, only after conducting this pilot study and Experiment 7 did we discover that another methodological approach would have been better suited to investigate how the different orientation profiles influence the landing position of the first fixation. This will be discussed in greater detail in Experiment 7, but the main point here is that even though the RoFF analysis revealed several significant main effects and interactions, these were nevertheless not particularly informative

as they were the by-products of the experimental design itself rather than the stimulus. However, the results are reported in Appendix A should the reader be interested.

8.1.2.1 Fixation Count

To ascertain if orientation contrast and orientation profiles affected the number of fixations on the stimuli, two ANOVAs were carried out. The first was a 2 (orientation contrast: high, and low) × 3 (orientation profile: Block, Blur, and Cornsweet) repeated measures ANOVA, and the second was a 2 (orientation contrast: high, and low) × 3 (orientation profile: Block, Blur, and Cornsweet) × 3 (region: centre, edge, and ground) repeated measures ANOVA. The first 2-way ANOVA was conducted using the absolute values of the fixation count data (Table 8.1, top), while the second 3-way ANOVA was conducted using the area-normalized fixation count data (Table 8.1, bottom. See Section 3.6.2 for area normalizing). The second ANOVA was specifically carried out to analyse the effect of orientation contrast and orientation profile on the different AOIs. As a result, the data had to be area normalized to account for the different sizes of the AOIs. For the area-normalized analysis (3-way ANOVA), only the region variable, and how it interacted with orientation contrast and orientation profile was explored. Main effects and an interaction of orientation contrast and orientation profile were not reported as they were explored in the 2-way ANOVA.

For main effects and interactions in which the assumption of sphericity was violated, the Greenhouse-Geisser ($\varepsilon<0.75$) or Huynd-Feldt ($\varepsilon>0.75$) correction was applied to the degrees of freedom and the p-values.

Table 8.1
ANOVA summary for fixation count data that was not area-normalized (top) and area-normalized (bottom). For analysis with respect to the different AOIs (Region variable), area-normalized data was calculated to take into account the different sizes of the AOIs.

	df	*F*	*p*
Fixation Count			
Contrast	1,23	13.005	**0.001**
Profile	2,46	4.015	**0.025**
Contrast × Profile	2,46	1.555	0.222
Area-Normalized Proportion of Fixation Count			
Region	1.03,23.76	249.396	**<0.001**
Contrast × Region	1.03,23.76	12.132	**0.002**
Profile × Region	1.83,41.98	6.774	**0.004**
Contrast × Profile × Region	1.92,44.19	3.822	0.063

The 2 × 3 repeated measures ANOVA (see Table 8.1 and Figure 8.2) revealed a significant main effect of orientation contrast, where Bonferroni corrected pairwise comparison showed that the low contrast stimuli (M=10.20) had more fixations than the high contrast stimuli (M=9.07) [p=0.001]. A main effect of Profile was also found, with Bonferroni corrected pairwise comparisons showing that participants made more fixations when viewing the Blur profile (M=9.90) compared to the Block profile (M=9.16) [p=0.048]. No other comparisons were significant. Finally, there was no interaction between orientation contrast and orientation profile.

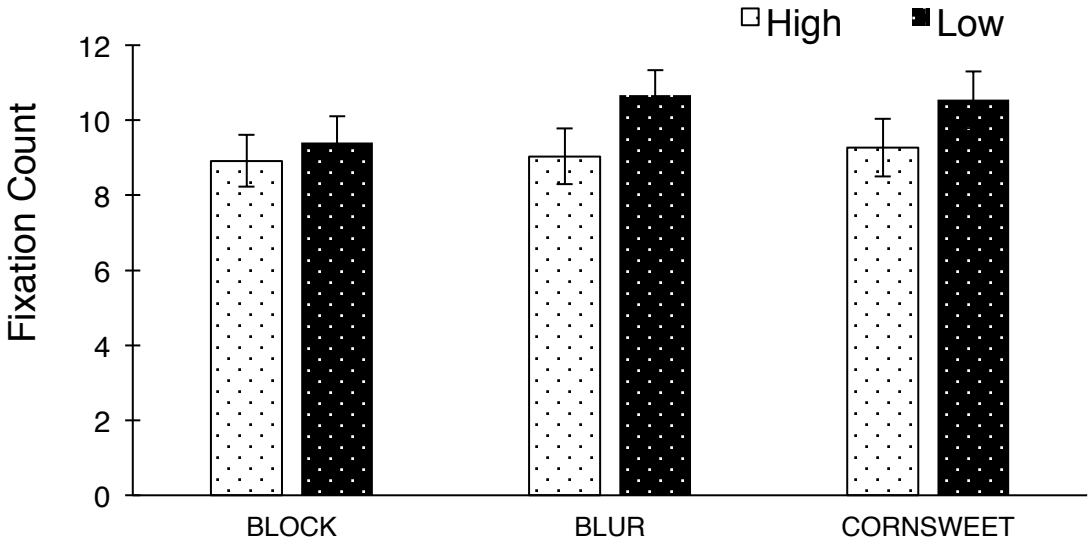


Figure 8.2. Average fixation count for the three orientation profiles across the different orientation contrast. Error bars represent the standard error of the mean.

The following analysis was done to determine how fixation count varied across the different AOIs. As a reminder, the data used to conduct this analysis was of the area-normalized proportion of fixations. That is, the proportion of fixation count to an AOI was divided by the proportion size of that AOI. The 2 × 3 × 3 repeated measures ANOVA (see Table 8.1 and Figure 8.3) revealed a significant main effect of region, which shows that the proportion of fixations to centre region (M=6.53) was the highest, followed by the edge region (M=1.43), and then the ground region (M=0.20) [for all Bonferroni corrected pairwise comparisons, p<0.001]. This implies that the centre region of a texture figure is fixated on most frequently, relative to its size.

Post hoc tests using the Bonferroni correction was used to analyse the significant contrast and region interaction. The comparisons showed that orientation contrast influenced the proportion of fixations to the centre (M_{High}=7.28; M_{Low}=5.83) [p=0.001] and ground region (M_{High}=0.12; M_{Low}=0.27) [p=0.001], but not the edge region (M_{High}=1.39; M_{Low}=1.47) [p=0.447]. That is, when orientation contrast increases, the proportion of fixations increases

to the centre region, and decreases to the ground region. This is an indication that participants were better at localising the figure patch when the contrast is higher.

Bonferroni corrected pairwise comparisons of the profile and region interaction show that the proportions of fixations to the edge region of the different profiles are not significantly different from each other (M_{Block}=1.38; M_{Blur}=1.28; $M_{Cornsweet}$=1.61) [for all comparisons, p>0.05]. However, for the centre region, the proportion of fixations to the Block profile (M=7.22) is greater than the Cornsweet profile (M=5.88) [p=0.007], and for the ground region, the proportion of fixations to the Block profile (M=0.13) is significantly lower than to the Blur (M=0.25) [p=0.001] and Cornsweet profile (M=0.22) [p=0.001].

Both these interactions (contrast × region and profile × region) seem to suggest that the proportion of fixations to the edge region is not influenced by orientation contrast or orientation profile. However, the proportion of fixations to the centre and ground region is affected by the orientation contrast and orientation profile, with an increase in fixation to the centre region and a decrease in fixation to the ground region when the stimuli segregate more easily as measured psychophysically i.e. Block profile.

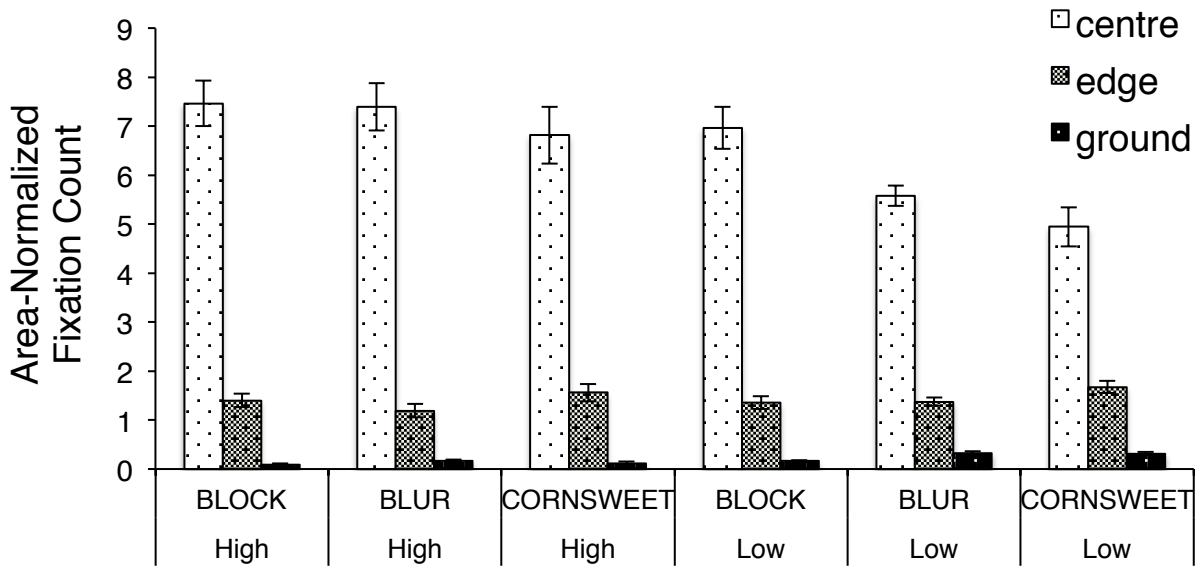


Figure 8.3. Average of all participants' data for the area-normalized proportion of fixation count for each orientation profile and orientation contrast. Error bars represent the standard error of the mean.

8.1.2.2 Summed Fixation Duration

To investigate the influence of orientation contrast and profile on the summed fixation durations to the different AOIs, area-normalized (see Section 3.6.2) summed fixation duration data was analysed with a 2 × 3 × 3 (orientation contrast × orientation profile × region) repeated measures ANOVA. Once again, only the region variable, and how it interacted with the orientation contrast and orientation profile was explored. Unlike the fixation count data described above, the absolute summed fixation duration data was not explored with a 2 × 3 (orientation contrast × orientation profile) ANOVA. This is because we fixed the stimulus duration to 5000ms, and the data gathered for summed fixation duration to the stimulus (as a whole as opposed to a particular region) would be approximately equal to each other.

For main effects and interactions in which the assumption of sphericity was violated, the Greenhouse-Geisser ($\varepsilon<0.75$) or Huynd-Feldt ($\varepsilon>0.75$) correction was applied to the degrees of freedom and the p-values.

Table 8.2
ANOVA summary for summed fixation duration analysis. Because the analysis is conducted with respect to the different AOIs (Region variable), area-normalized data was calculated to take into account the different sizes of the AOIs.

	df	*F*	*p*
Area-Normalized Proportion of Summed Fixation Duration			
Region	1.02,23.52	235.787	**<0.001**
Contrast × Region	1.04,23.88	12.758	**0.001**
Profile × Region	1.99,45.83	4.916	**0.001**
Contrast × Profile × Region	2.10,48.31	5.306	0.053

The data was area-normalized by dividing the proportion of summed fixation duration to an AOI by the proportion size of the AOI. This analysis allowed us to determine how the summed fixation duration varied across the different AOIs. A 2 × 3 × 3 repeated measures ANOVA (see Table 8.2 and Figure 8.4) revealed a significant main effect of region, in which Bonferroni corrected pairwise comparisons show that the proportion of summed fixation duration to centre region (M=7.21) was the highest, followed by the edge region (M=1.28), and then the ground region (M=0.16) [for all comparisons, $p<0.001$]. This implies that the centre region of a texture figure is fixated on for a longer time compared to the edge region.

Post hoc tests using the Bonferroni correction was used to analyse the significant contrast and region interaction. The comparisons show that orientation contrast influenced the proportion of summed fixation duration to the centre (M_{High}=7.87; M_{Low}=6.55) [p=0.001] and ground region (M_{High}=0.09; M_{Low}=0.22) [$p<0.001$], but not the edge region (M_{High}=1.23; M_{Low}=1.34) [p=0.270]. This shows that as orientation contrast increases, the proportion of summed fixation duration to the centre region increases, while it decreases to the ground region. This

once again indicates that participants were better at localising the figure patch when the figure was more salient.

Bonferroni corrected pairwise comparisons of the profile and region interaction show that the proportion of summed fixation duration to the edge region of the different profiles is not significantly different (M_{Block}=1.24; M_{Blur}=1.12; $M_{Cornsweet}$=1.50) [for all comparisons, p>0.05]. However, for the centre region, the proportion of summed fixation duration to the Block profile (M=7.74) is greater than the Cornsweet profile (M=6.50) [p=0.019]. For the ground region, the proportion of summed fixation duration to the Block profile (M=0.11) is significantly lower than the Blur (M=0.19) [p=0.004] and Cornsweet profile (M=0.18) [p=0.003].

These results (contrast × region interaction and profile × region interaction) suggest that neither orientation contrast or orientation profile influences the proportion of summed fixation duration to the edge region. However, it does influence the proportion of summed fixation duration to centre and ground region. That is, when a texture figure segregates easily, more time is spent on the centre region, and less to the ground region. This is the same pattern of results we find with the area-normalized proportion of fixation count.

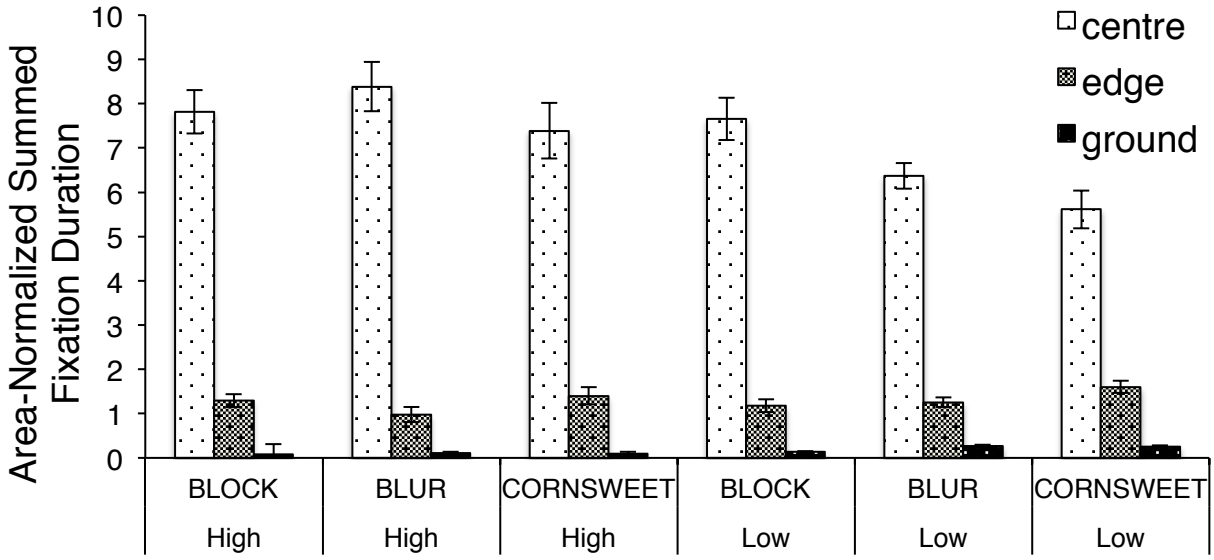


Figure 8.4. Average of all participants' data for the area-normalized proportion of summed fixation duration for each orientation profile and orientation contrast. Error bars represent the standard error of the mean.

8.2 Experiment 7: Task to Locate Centre or Edge of Figure

In the pilot study, participants' eye movements were tracked while they located a figure patch that segregated from the background due to change in orientation. We found that the different orientation profiles and orientation contrast did show some differences in eye movement patterns. However, participants were only required to look for the figure, and there was no objective measure of whether or not the participants were actually able to detect the target figure. Therefore, it is possible that on some trials, especially with the low contrast stimuli, participants were not in fact able locate the figure. In those instances, the eye movements measured would not have been a result of the figure per se, but due to participants scanning the texture looking for the figure.

Hence, in the current experiment, participants' eye movements were tracked while they performed two tasks. The two experimental tasks were the Centre Task and the Edge Task. In

the Centre Task, participants were required to indicate via a mouse-click the location of the centre of the figure, whereas for the Edge Task they had to indicate via a mouse-click the location of the edge of the figure. This study gives us two additional pieces of information that the previous study did not: 1) are participants able to detect the presence of a figure? 2) are participants able to localize the shape of the figure (i.e. segment the figure, not just detect it)?

In this study, we have measured behavioural data in terms of accuracy in detecting the figure centre and edge, and also the time it takes to detect the figure centre and edge. We hypothesise for both behavioural measures that performance will be better when orientation contrast is high. Moreover, performance will be better for the Block profile, followed by the Cornsweet, then the Blur. In fact, this pattern of results is also expected for the eye-movement measures of fixation count and summed fixation duration, whereby we expect that stimuli that are less detectable will have a greater number of fixations, and higher summed fixation durations.

To investigate where participants look in a texture stimulus, and how this is affected by orientation contrast and the different orientation profiles, eye movement measures were again analysed in terms of Areas of Interests (AOIs). As mentioned in the previous study, a different experimental design from that used in this study and the pilot study would have been better suited to investigate how the orientation profiles affect the location of the first fixation. However, it still does inform about the patterns of eye movements elicited when searching for an embedded figure in texture stimuli.

We hypothesised that regions of a texture figure that produces the most information regarding orientation contrast between figure and ground would be fixated on more often and for longer times. We further hypothesise that the task demand would introduce top-down factors that

would influence eye movements to the different regions. Therefore, we would expect that the number of fixations and summed fixation duration would be higher to the edge region for the Edge Task, but higher to the centre region for the Centre task.

8.2.1 Methods

8.2.1.1 Participants

Twenty-four participants (M_{age} = 19.9, SD_{age} = 2.8) were recruited from the University of Nottingham Malaysia. All participants had normal, or corrected-to-normal vision.

8.2.1.2 Apparatus & Display

The apparatus setup and display were the same as the pilot study (see section 8.1.1.2).

8.2.1.3 Stimuli

The experimental stimuli used were identical to the pilot study (see section 8.1.1.3). However, in addition to the 30 figure-ground texture stimuli, eight uniform orientation stimuli were included (Figure 8.5) as catch trials. These stimuli contain no target figures as they consist of line elements with the same mean orientation throughout the 32×32 grid. Both tasks, which are described below in the procedure, used the same set of stimuli, presented in randomized order.

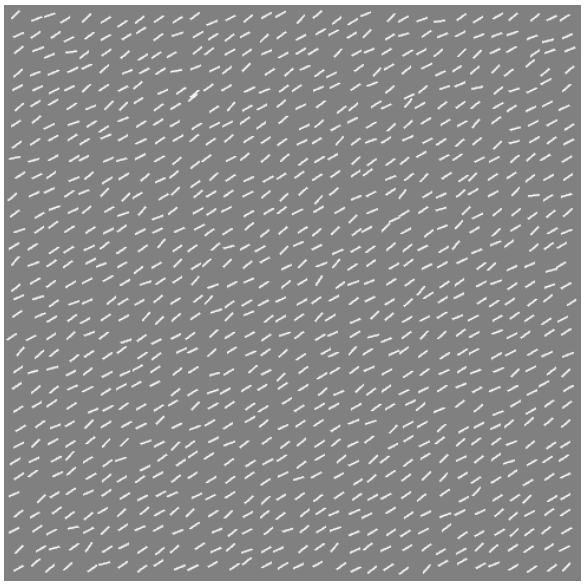


Figure 8.5. Stimuli in which there is no segregated patch as the orientation of the line elements are all uniform. Orientation and positional jitter were the same as the figure-ground texture.

8.2.1.4 Procedure

Prior to starting the experiment, the experimenter showed the participants examples of stimuli that would appear in the experimental session. These examples were different to those that the participants were actually shown in the experimental session. The participants were shown these examples of stimuli on the eye tracker monitor, and the experimenter explicitly pointed out to them the location of the centre (middle of figure) and edge region (along the border of the rectangle) of the figure. Participants were also told that all of the four edge borders represent the edge of the figure.

For both the tasks, a fixation cross was presented on the left side of the screen for 1000ms, followed by the appearance of the texture stimuli. Participants were instructed to use the left mouse button to indicate the edge (Edge Task) or centre (Centre Task) of the figure as fast and as accurately as possible. If participants did not detect a figure within the texture, they

were instructed to use the right mouse button anywhere beyond the region of the 32X32 grid of line elements. The mouse press ended the trial.

Specifically, participants were instructed: *"You are to determine whether or not a smaller figure is embedded in the larger square. When you do identify the small figure, use the mouse to click on the centre region"* Participants were also instructed to: *"Try and respond as fast as possible, as well as be accurate"*.

Before each task, instructions were repeated to the participants, and they performed a practice session before moving on to the experimental session. The order in which the tasks were presented to the participants was counterbalanced i.e. half the participants performed the Edge Task first, while the other half performed the Centre Task first.

8.2.2 Results: Behavioural Responses

8.2.2.1 Accuracy

Accuracy scores are plotted as a percentage of correct responses, which refers to the participants' ability to identify the edge (Edge Task) or centre (Centre Task) of the figure correctly. Trials in which participants did not report seeing a figure when a figure was present were also classified as incorrect responses.

Edge Task. A 2 (orientation contrast: high and low) × 3 (orientation profile: Block, Blur, and Cornsweet) repeated measures ANOVA (see Figure 8.6) examining accuracy of performance to detect the edge of the figure showed a significant main effect of orientation contrast [$F(1,23) = 59.147$, $p<0.001$]: where Bonferroni corrected pairwise comparison shows that participants were more accurate at identifying edges of figures with high orientation contrast (M=97.2%) compared to low orientation contrast (M= 73.3%). A main effect of Profile was

also observed [$F(2,46)=28.401$, $p<0.001$]: where performance was most accurate for the Block profile (95%), followed by performance for the Cornsweet profile (84.6%), and the Blur profile having the least accurate performances (76.3%) [where Bonferroni corrected pairwise comparisons reveal that $p<0.001$ for Block vs. Blur and Block vs. Cornsweet, $p=0.024$ for Blur vs. Cornsweet].

Finally, there was a significant interaction between Contrast and Profile [$F(2,46)= 9.514$, $p<0.001$]. Pairwise comparisons using the Bonferroni correction show that for the high contrast stimuli, performance was worse for the Blur profile ($M=91.7\%$) compared to the Block profile ($M=100\%$) [$p=0.006$] and Cornsweet profile ($M=100\%$) [$p=0.006$]. However, for the low contrast stimuli, performance was worse for both the Blur ($M=60.8\%$) [$p<0.001$] and Cornsweet ($M=69.2\%$) [$p=0.001$] profile compared to the Block profile ($M=90\%$). This suggests that the Block profile, which has information regarding orientation contrast at both figure edge and centre, is best for localising the edge of the figure, especially when saliency is low.

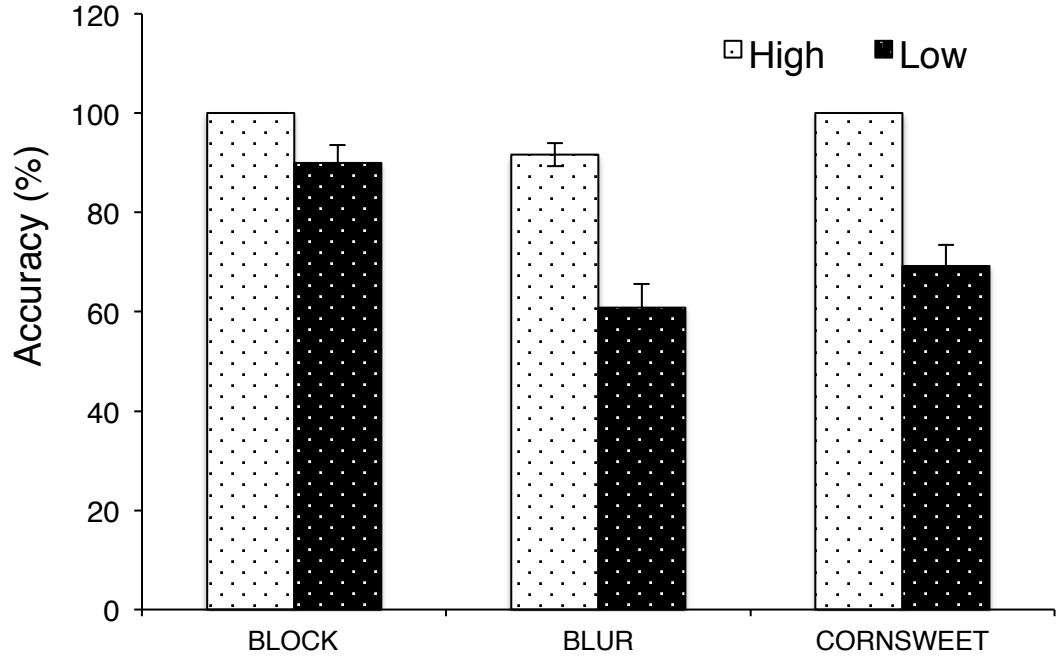


Figure 8.6. Accuracy results for the Edge Task. Error bars represent the standard error of the mean. Where no error bars are seen, there was no error estimate associated with the mean value i.e. absolutely no deviation between participants' response.

Centre Task. A 2 (orientation contrast: high and low) × 3 (orientation profile: Block, Blur, and Cornsweet) repeated measures ANOVA (see Figure 8.7) examining accuracy of performance to detect the centre of the figure was also carried out. A significant main effect of Contrast was found [$F(1,23) = 58.030$, $p<0.001$]: in which Bonferroni corrected pairwise comparison show that participants were more accurate at identifying the centre of figures with high orientation contrast (M=99.4%) compared to low orientation contrast (M= 80.3%). A main effect of Profile was also observed [$F(2,46)= 15.952$, $p<0.001$]: where Bonferroni corrected pairwise comparisons reveal that performance for the Cornsweet profile (85.4%) [p=0.001] and the Blur profile (87.1%) [$p<0.001$] were lower compared to the performance of the Block profile (99.1%).

Finally, there was a significant interaction between Contrast and Profile [$F(2,46)= 13.591$, $p<0.001$], where performance for the 3 profiles were not significantly different from each other for the high contrast stimuli (M_{Block}=100%, M_{Blur}=98.3%, $M_{Cornsweet}$=100%) [$p>0.05$ for all pairwise comparisons using the Bonferroni correction]. However, for the low contrast stimuli, performance was worse for both the Blur (M=75.8%) [$p<0.001$] and Cornsweet profile (M=70.8%) [p=0.003] compared to the Block profile (94.2%). Thus, when saliency is low, the Block profile shows best performance at localising the centre of the figure.

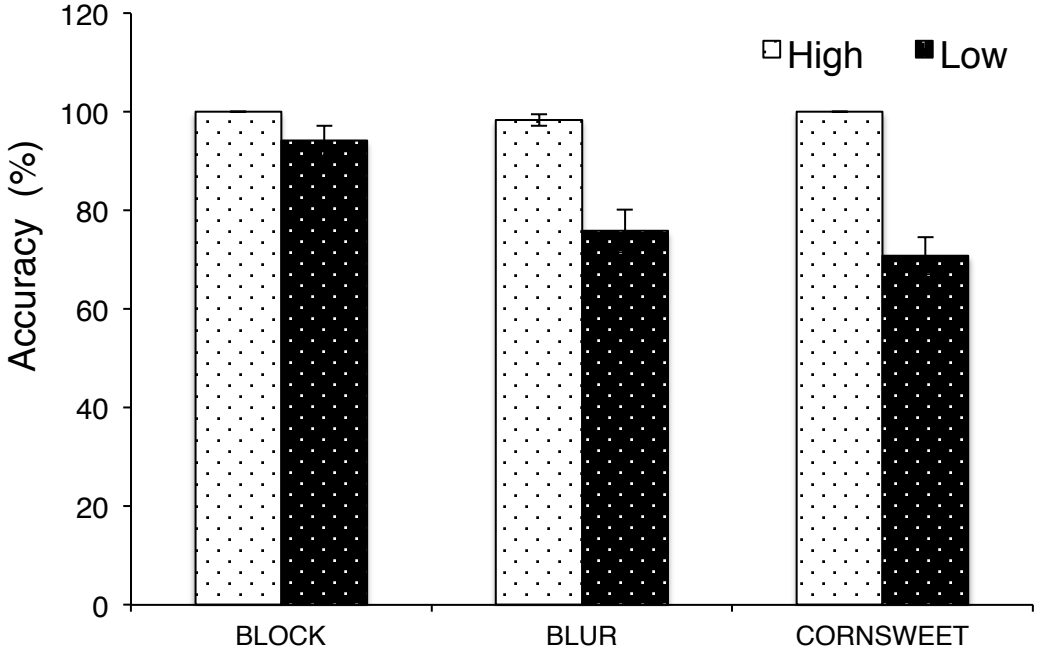


Figure 8.7. Accuracy results for the Centre Task. Error bars represent the standard error of the mean. Where no error bars are seen, there was no error estimate associated with the mean value i.e. absolutely no deviation between participants' response.

8.2.2.2 Time to Mouse-Click

Time to mouse-click is the time taken for participants to make a response to the edge (Edge Task) or the centre (Centre Task) of the figure after stimulus onset. Trials with incorrect responses were removed from the analysis beforehand.

Edge Task. A 2 (orientation contrast: high and low) × 3 (orientation profile: Block, Blur, and Cornsweet) repeated measures ANOVA (see Figure 8.8) was used to analyse the time taken to make a response to the edge of the figure. A significant main effect of Contrast [$F(1,23) = 22.694$, $p<0.001$] was found, where Bonferroni corrected pairwise comparison shows that participants were faster at identifying the edge of a figure with high orientation contrast (M=2480ms) compared to low orientation contrast (M= 4510ms). A main effect of Profile was also observed [$F(2,46)= 6.735$, p=0.003], where performance was faster for the Block profile (M=2890ms) compared to the Blur profile (M=4280ms) [p=0.003]. For the Cornsweet profile (M=3320ms), reaction times were not different compared to the Block [p=0.46] and Blur [p=0.17] profiles. There was no significant interaction between Contrast and Profile [$F(2,46)= 1.279$, p=0.288].

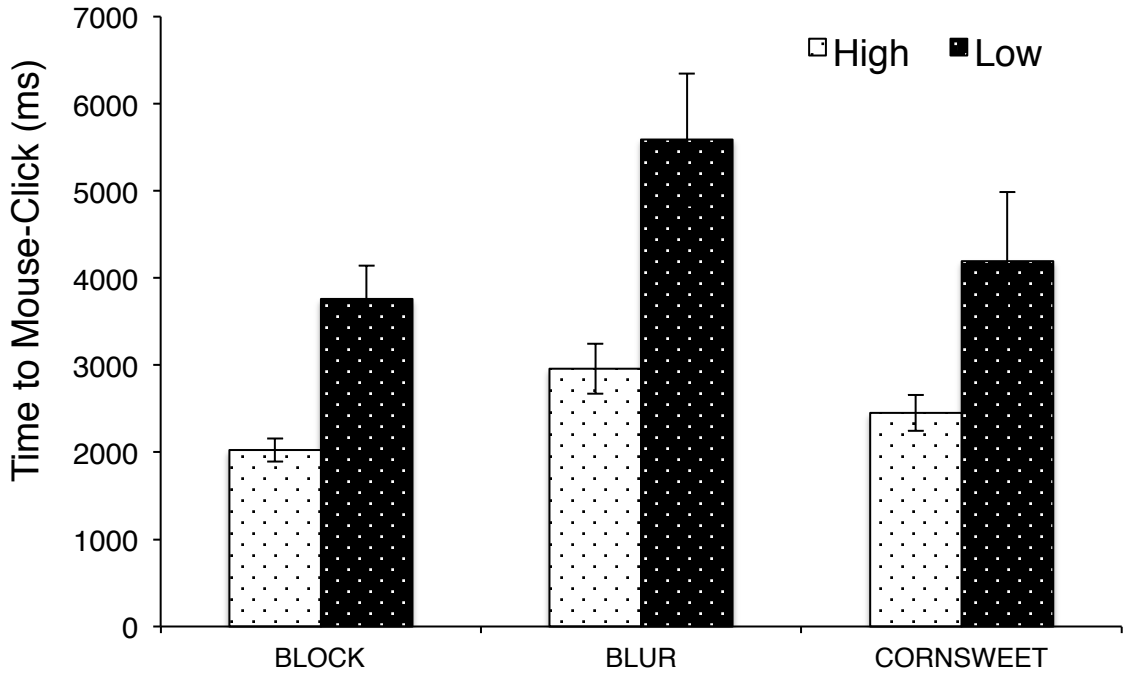


Figure 8.8. Time taken to make a mouse-click on the edge of the figure. Error bars represent the standard error of the mean.

Centre Task. A 2 (orientation contrast: high and low) × 3 (orientation profile: Block, Blur, and Cornsweet) repeated measures ANOVA (see Figure 8.9) was used to analyse the time taken to make a response to the centre of the figure. The results showed a significant main effect of Contrast [$F(1,23) = 17.483$, $p<0.001$]: where pairwise comparisons using the Bonferroni correction show that participants were faster at identifying the centre of a figure with high orientation contrast (M=2220ms) compared to low orientation contrast (M=4430ms). A main effect of Profile was also found [$F(2,46)= 6.242$, p=0.004], where performance was faster for the Block profile (M=3040ms) compared to the Blur profile (M=4010ms) [$p<0.001$]. For the Cornsweet profile (M=2930ms), reaction times were not faster compared to the Block [p=1.000] and Blur [p=0.069] profiles. There was no significant interaction between Contrast and Profile [$F(2,46)= 2.470$, p=0.096].

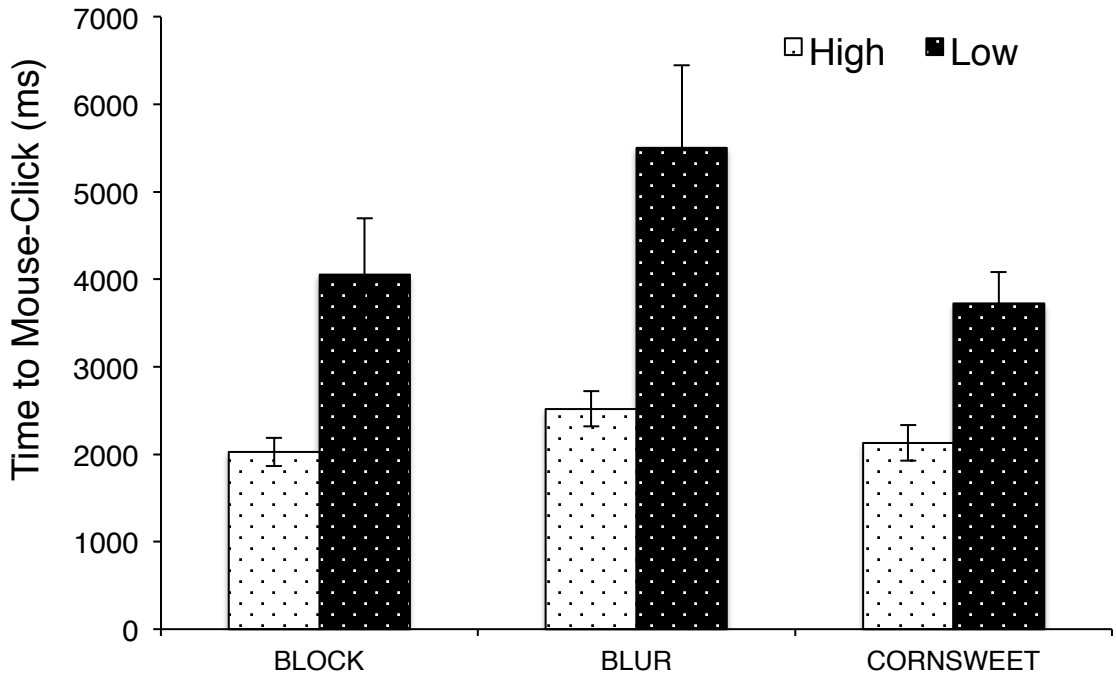


Figure 8.9. Time taken to make a mouse-click on the centre of the figure. Error bars represent the standard error of the mean.

8.2.3 Results: Eye Tracking Measures

Before analysing the eye tracking data, incorrect trials were first removed. These were trials in which participants failed to correctly identify the edge/centre or failed to detect the presence of the figure. These trials were excluded from all the eye tracking measures – position of first fixation, region of first fixation (RoFF), saccade latency, fixation count, and summed fixation duration.

8.2.3.1 Position of First Fixation

We analysed the position of first fixations in terms of their xy-coordinates. That is, for each of the 30 stimuli used, the landing positions of the first fixation of all participants were plotted and analysed. However, it was only after conducting this analysis that we discovered that the methodological approach used in the current study was not suited to investigate how orientation profile and orientation contrast influences the landing position of the first fixation. As a result, even though this was one of the last analysis conducted for this study, it will be described here first before all the other analysis to illustrate to the reader why the RoFF analysis is not particularly informative.

Figure 8.10 shows the landing positions of the first fixation for two stimuli (see Appendix B for the plots of all stimuli). For each plot, we calculated the centroid, represented by an × in the plots, which is the mean landing position of first fixation for all participants. Both the plots shown in Figure 8.10 are responses to the high orientation contrast Blur profile when performing the Edge Task. Despite the fact that the figure patch in both stimuli is positioned in different places from each other, the landing positions of the first fixation for all participants are clustered in the same region.

Subsequently, we plotted the centroids for all the 30 stimuli separately for the Edge and Centre Task (see Figure 8.11). By visual inspection, we can see that the centroids are all clustered together regardless of the stimulus profile, orientation contrast, or task. We then calculated the average position of all the centroids for each task separately i.e. the centroid of all the centroid points, which we will refer to as the global centroid. In terms of pixel coordinates, the global centroids were [425,367] and [442,374] respectively for the Edge and Centre Task. Both these points are very close to each other (a distance of 30'), implying that the task does not affect the position of first fixation. In terms of percentage distance of the global centroid relative to the centre of the texture grid, these were 83% and 86% respectively. Thus, it suggests that the participants' first fixation was being directed to the midpoint of the entire texture grid, with a shortfall of ≈ 15.5%.

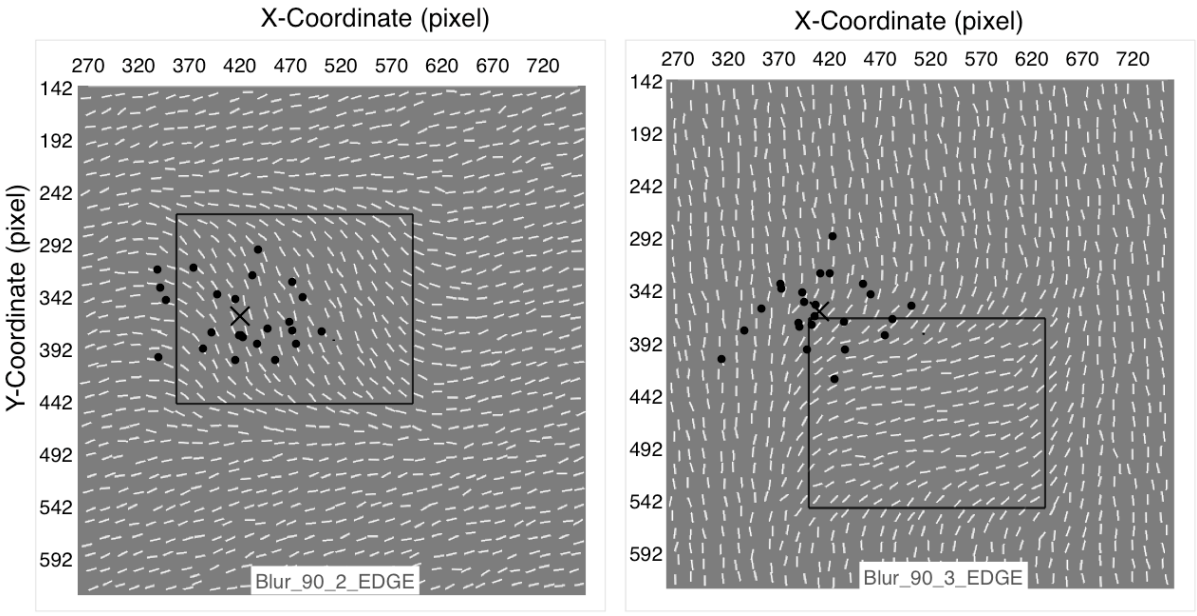


Figure 8.10. Landing positions of the first fixation in response to the Edge Task for two examples of stimuli with high orientation contrast Blur profile. Each black dot is one participants landing position, and the × represents the centroid, which is the mean position of all the participants' landing position. The centroid is positioned at [420,361] for the image on the left and [409,362] for the image on the right, which is a difference of 11 pixels/17'. The solid line represents the border of the figure patch.

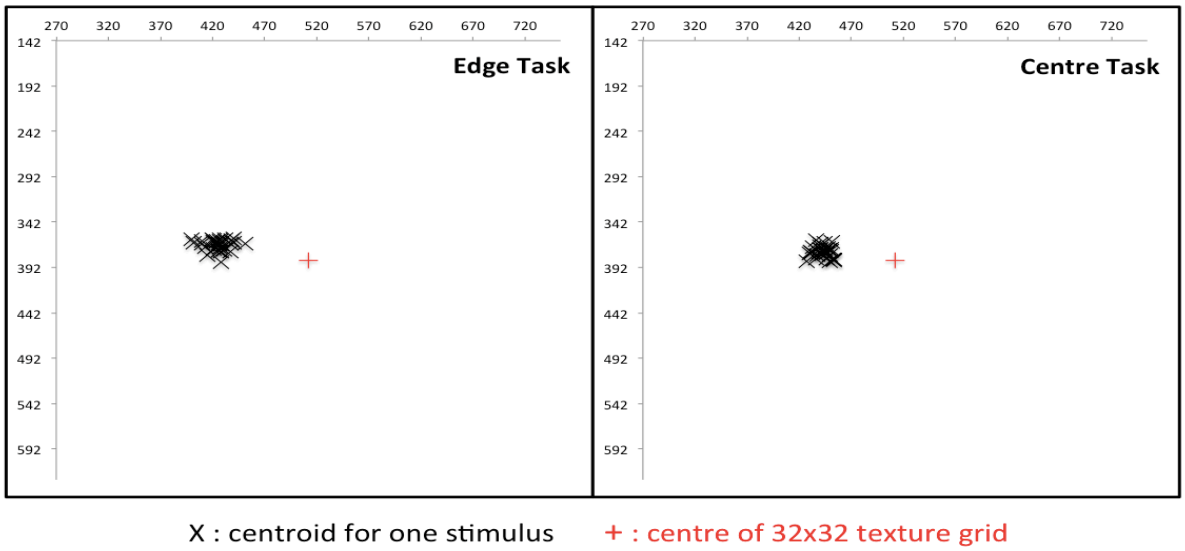


Figure 8.11. The graph shows the plots of centroids (×) for each of the 30 stimuli for the Edge (left figure) and Centre (right figure) tasks separately. The global centroid for the Edge Task was [425,367] whereas it was [442,374] for the Centre Task. The XY-axes represent the position of the texture grid in pixels.

As well as calculating the centroids of the location of first fixation, we also analysed the standard deviation of the location of first fixation for each stimulus. If the type of stimuli affects where participants look, the standard deviation should vary between conditions as the profile, task, and contrast are also affecting the fixation position. On the other hand, if the standard deviation does not vary between the conditions, it would once again suggest that participants' location of first fixation is slightly to the left of the middle of the texture grid, with some degree of uniform variation.

For the analysis, the standard deviation for each of the 30 stimuli was calculated. This was obtained by calculating the standard deviation of the 24 participants' data points for each stimulus. A 2 (task: edge and centre) × 3 (profile: Block, Blur, and Cornsweet) × 2 (orientation contrast: high and low) repeated measures ANOVA was carried out. Table 8.3 summarises the results of the ANOVA, which shows no significant main effects or

interactions were found. This further supports the notion that the position of first fixation was not influenced by the stimulus profile, orientation contrast, or task.

Table 8.3
Summary of ANOVA analysis on standard deviation data.

	df	*F*	*p*
Task	1,4	0.901	0.396
Profile	2,8	0.537	0.604
Contrast	1,4	5.386	0.081
Task × Profile	2,8	2.309	0.162
Task × Contrast	1,4	1.545	0.282
Profile × Contrast	2,8	1.480	0.284
Task × Profile × Contrast	2,8	0.787	0.488

After analysing the deviation between the participants' position of first fixation, we further analysed the centroid results by making two different comparisons. For the first comparison, we calculated the distance between the centroids and the middle of the figure patch. If the participants were actually making a first fixation based on the position of the figure patch, the distance between the centroids and the middle of the figure patch should be small. For the second comparison, the distance between the centroids and the global centroid was calculated. The two comparisons were graphed as frequency distributions (see Figure 8.12).

The distance between the centroid and the middle of the figure patch (Figure 8.12, left column) ranges from 20-140 pixels (mean = 84.9 pixels or 2.2°) for the Centre Task and 40-160 pixels (mean = 99.0 pixels or 2.5°) for the Edge Task. As mentioned above, if the overall position of first fixation were influenced by the position of the figure, the distance between the centroids and the middle of the figure patch should be small. Our results here show that the distance is actually large, suggesting that the position of first fixation was not affected by the figure position. However, the distance between the centroids and the global centroid was much smaller (Figure 8.12, right column). For the Centre Task, it ranged between 0.67 – 18

pixels (mean = 8.5 pixels or 0.2°), and for the Edge Task, it ranged between 0.64 – 24 pixels (mean = 10.6 pixels or 0.3°). This indicates that participants were making first fixations to a similar region, which was just to the left of the centre point of the texture grid.

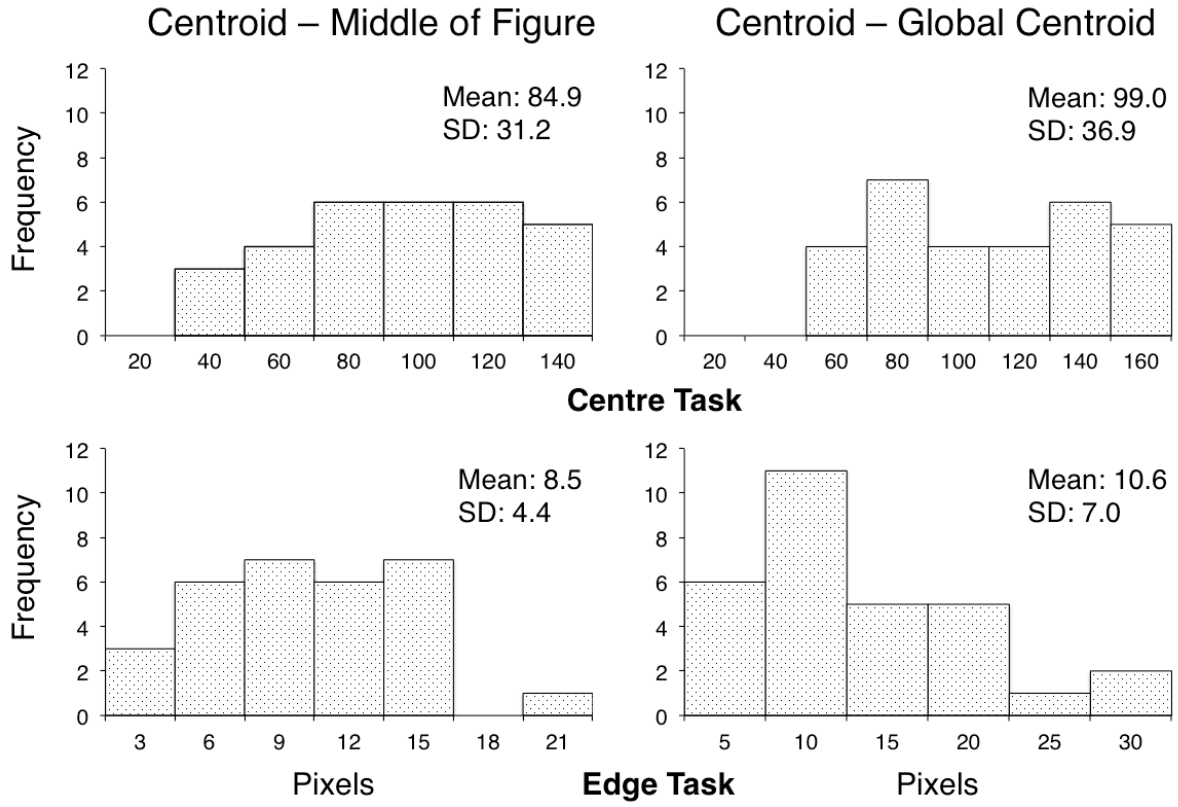


Figure 8.12. Frequency distributions of the distance between the centroids and the middle of the figure patch (left) and centroids and the global centroid (right) for the centre (top) and edge (bottom) tasks.

From the various analyses conducted, we posit that the position of first fixation was not being influenced by location of the figure patch, nor was it influenced by the stimulus profile, orientation contrast, or even the task. It does however raise the question of why the RoFF analysis indicates a difference in the proportion of first fixations to each AOI for the various stimulus conditions (see Appendix C for RoFF analysis). To investigate this, we re-examined the plots of the position of first fixation for each stimuli created (Appendix B).

We found that even though the figure was randomly positioned within the texture grid, there were only 5 examples per condition, and there were differences in figure position between each condition. For example, for the Block profile with low orientation contrast, 3 out of the 5 stimuli had the left edge of the figure positioned where the centroid position is. However, for the Block high contrast condition, only 1 out of the 5 stimuli had the left edge of the figure positioned where the centroid position is. Therefore, with the RoFF analysis, it would appear as though there were more frequent first fixations to the Edge region for the Block low contrast condition than the Block high contrast condition. However, that has nothing to do with the condition itself, but the position of the figure in relation to the centroid position. As a consequence of this, the outcome of the RoFF data analysis has more to do with the particular stimulus examples used rather than the mechanisms driving eye movements.

8.2.3.2 Saccade Latency

The saccade latency measures the time it takes for a participant to initiate movement to the first fixation. The time starts when the stimulus first appears on screen, and ends when the saccade is initiated for the first time. The saccade duration on average was ≈50ms, which was the time between the first saccade initiation and the first fixation. The saccade latency was analysed with a 2 (orientation contrast: high, and low) × 3 (orientation profile: Block, Blur, and Cornsweet) repeated measures ANOVA for both the centre and Edge Task. The analysis revealed that the orientation contrast and profile of the figure did not influence saccade latency (see results below). Subsequently, the data of all participants from all trials were grouped together according to task, and were plot as a frequency distribution (Figure 8.13). The shape of the distribution is similar for both tasks. These results suggest that saccade latency was not influenced by the stimuli or the task, and on average participants require approximately 250ms to initiate a saccade.

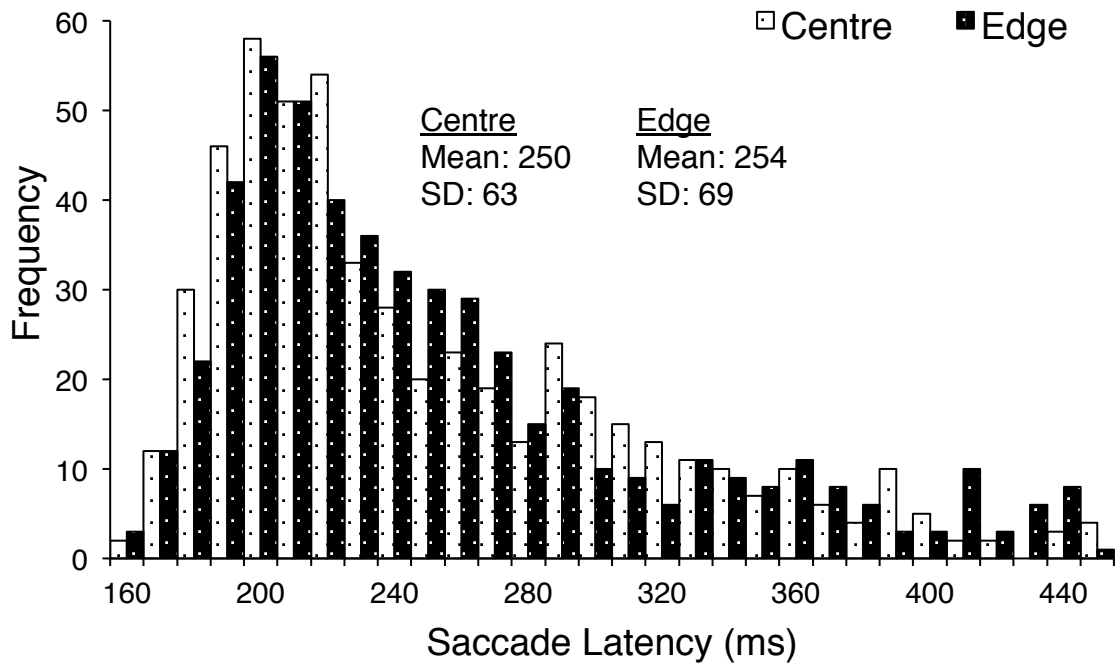


Figure 8.13. A frequency distribution plot of saccade latencies. The histogram plot is of pooled latencies of saccades from all participants. Note how the distribution of both tasks is very similar.

Edge Task. The 2 × 3 ANOVA (see Figure 8.14) revealed that there was no main effect of contrast, [$F(1,21) = 0.351$, $p=0.560$] or profile [$F(2,42) = 2.163$, $p=0.128$]. There was no significant interaction either [$F(2,42) = 3.075$, $p=0.057$]. Therefore, neither the orientation contrast or orientation profile influenced the time taken to initiate a saccade.

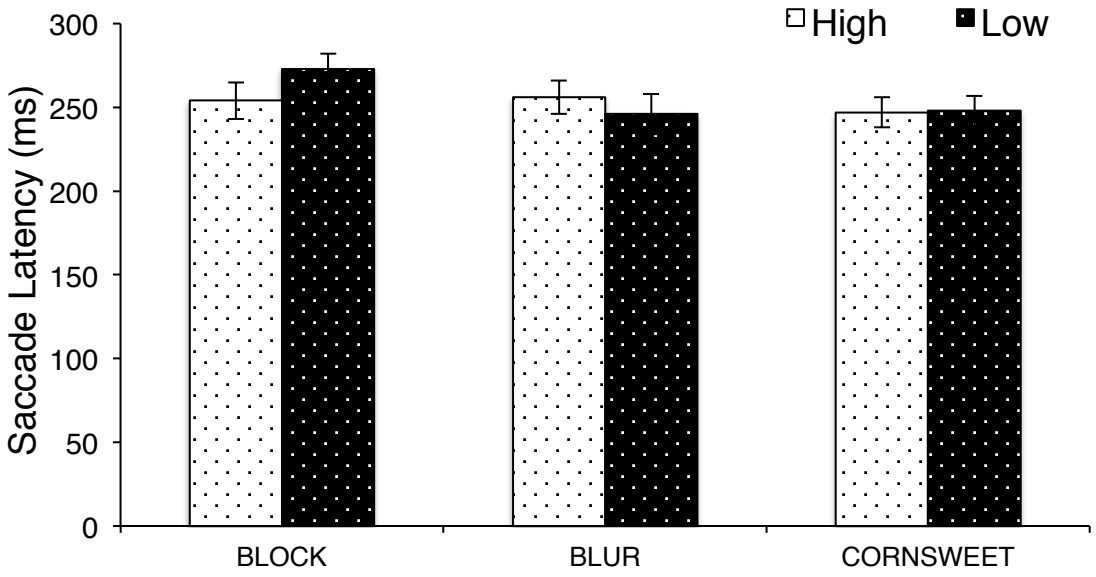


Figure 8.14. Average saccade latency for the three orientation profiles across the different orientation contrast in response to the Edge Task. Error bars represent the standard error of the mean.

Centre Task. The 2 × 3 ANOVA (see Figure 8.15) revealed that there was no main effect of contrast, [$F(1,23) = 0.555$, $p=0.464$] or profile [$F(2,46) = 0.945$, $p=0.396$]. There was no significant interaction either [$F(2,46) = 1.873$, $p=0.165$]. Once again, the results indicate that orientation contrast and orientation profile did not influence the time taken to initiate a saccade.

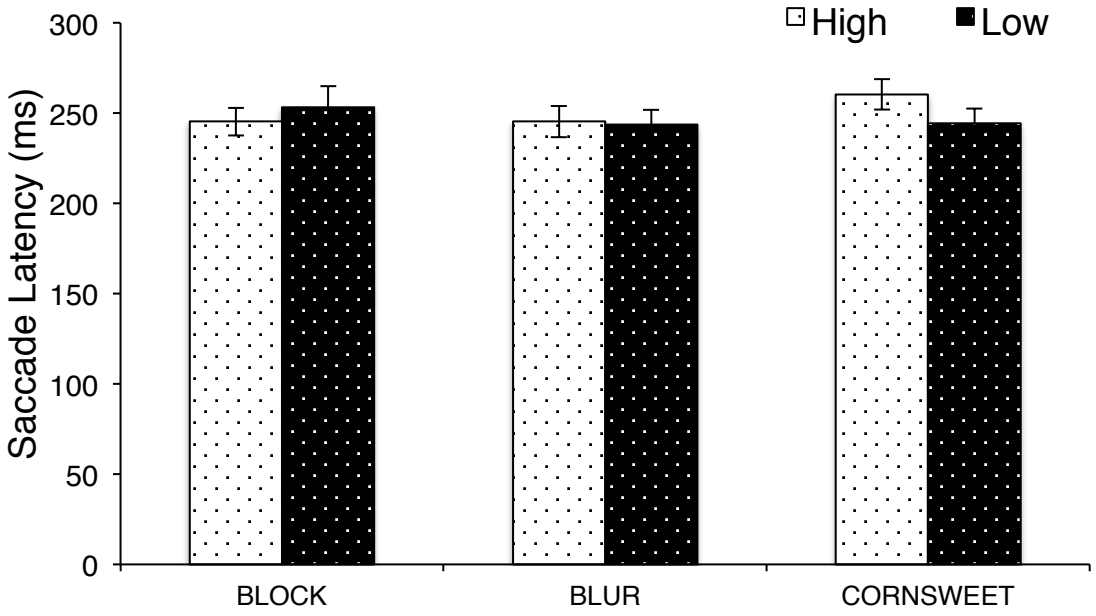


Figure 8.15. Average saccade latency for the three orientation profiles across the different orientation contrast in response to the Centre Task. Error bars represent the standard error of the mean.

8.2.3.3 Fixation Count

For both the Edge and Centre Task, the fixation count data was analysed in two ways. First, a 2 × 3 (orientation contrast × orientation profile) ANOVA was conducted using the absolute values of fixation count data. This was conducted to investigate whether or not orientation contrast and orientation profiles influenced the number of fixations on a stimulus. Subsequently, a 2 × 3 × 3 (orientation contrast × orientation profile × region) ANOVA was conducted. This ANOVA was carried out to analyse the effect of orientation contrast and orientation profile on the different AOIs, hence the inclusion of the region variable. However, due to the different sizes of the AOIs, area normalized fixation count data was used in this 3-way ANOVA (see section 3.6.2). For the area-normalized analysis, only the region variable, and how it interacted with the orientation contrast and orientation profile were explored.

Main effects and an interaction of orientation contrast and orientation profile were not reported as they were explored in the 2-way analysis.

For main effects and interactions in which the assumption of sphericity was violated, the Greenhouse-Geisser ($\varepsilon<0.75$) or Huynd-Feldt ($\varepsilon>0.75$) correction was applied to the degrees of freedom and the p-values.

Edge Task. The 2 × 3 repeated measures ANOVA (see Table 8.4 and Figure 8.16) revealed a significant main effect of orientation contrast, in which Bonferroni corrected pairwise comparison shows that the low contrast stimuli (M=11.32) had more fixations than the high contrast stimuli (M=6.24). A main effect of Profile was observed, where participants made more fixations to the Blur profile (M=11.17) compared to the Block (M=7.04) [$p<0.001$] and Cornsweet profile (M=8.17) [p=0.005]. These results show that more salient targets require fewer fixations.

Finally, there was a significant interaction between Contrast and Profile. Post hoc tests using the Bonferroni correction show that when orientation contrast was high, the number of fixations to the Blur (M=7.26) [p=0.004] and Cornsweet (M=6.62) [p=0.015] profiles was higher than to the Block profile (M=4.84), but when orientation contrast was low, only the Blur profile (M=15.08) had higher number of fixations compared to the Block (M=9.23) [$p<0.001$] and Cornsweet (M=9.66) [p=0.004] profiles.

Table 8.4
ANOVA summary for fixation count data that was not area normalized (top) and area normalized (bottom) for the Edge Task. For analysis with respect to the different AOIs (Region variable), area-normalized data was calculated to take into account the different sizes of the AOIs.

	df	*F*	*p*
Fixation Count			
Contrast	1,21	36.325	**<0.001**
Profile	2,42	16.429	**<0.001**
Contrast × Profile	2,42	6.066	**0.005**
Area-Normalized Proportion of Fixation Count			
Region	1.03,23.72	109.634	**<0.001**
Contrast × Region	1.06,24.37	4.468	**0.043**
Profile × Region	2.15,49.39	4.845	**0.001**
Contrast × Profile × Region	2.13,49.01	0.390	0.815

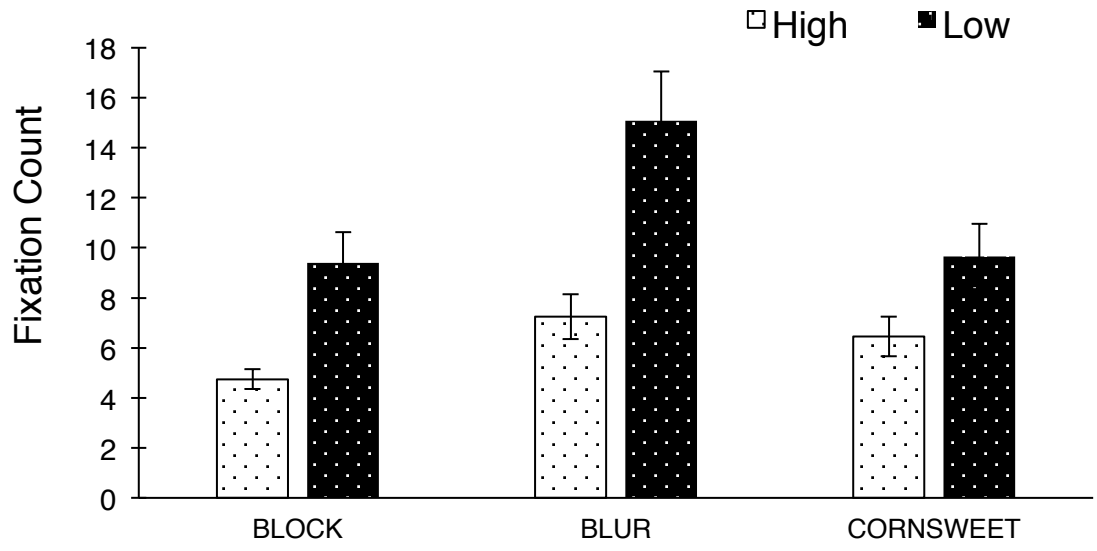


Figure 8.16. Average fixation count for the three orientation profiles across the different orientation contrasts in response to the Edge Task. Error bars represent the standard error of the mean.

The following analysis was done to determine how fixation count varied across the different AOIs. As a reminder, the data used to conduct this analysis was of the area-normalized proportion of fixations. That is the proportion of fixation count to an AOI was divided by the

proportion size of that AOI. The 2 × 3 × 3 repeated measures ANOVA (see Table 8.4 and Figure 8.17) revealed a significant main effect of region, which shows that the proportion of fixations to centre region (M=4.54) was the highest, followed by the edge region (M=2.45), and then the ground region (M=0.11) [for all Bonferroni corrected pairwise comparisons, p<0.001]. This is an unexpected outcome as we initially hypothesised that the proportion of fixations to the edge region would be greater than the centre region when performing the Edge Task. However, the results show that irrespective of the task demand, the centre region is fixated on more frequently.

Post hoc tests using the Bonferroni correction was used to analyse the significant contrast and region interaction. The comparisons show that orientation contrast influenced the proportion of fixations to the centre (M_{High}=4.30; M_{Low}=4.79) [p=0.016] and edge region (M_{High}=2.66; M_{Low}=2.24) [p=0.004], but not the ground region (M_{High}=0.10; M_{Low}=0.13) [p=0.109]. Therefore, as orientation contrast increases, the proportion of fixations decreases to the centre region, but increases to the edge region. This suggests that as the target figure becomes more detectable, the proportion of fixations to edge region increases.

Bonferroni corrected pairwise comparisons of the profile and region interaction show that the proportion of fixations to centre region is not influenced by the profiles (M_{Block}=4.59; M_{Blur}=4.78; $M_{Cornsweet}$=4.26) [for all comparisons, p>0.05]. For the edge region, the proportion of fixations to the Blur profile (M=2.07) is lower than the Block (M=2.66) [p<0.001] and Cornsweet (M=2.62) [p=0.001] profile, while for the ground region, the proportion of fixations to the Blur profile (M=0.16) is the highest, followed by the Cornsweet profile (M=0.11), and least for the Block profile (M=0.07) [for all comparisons, p>0.05].

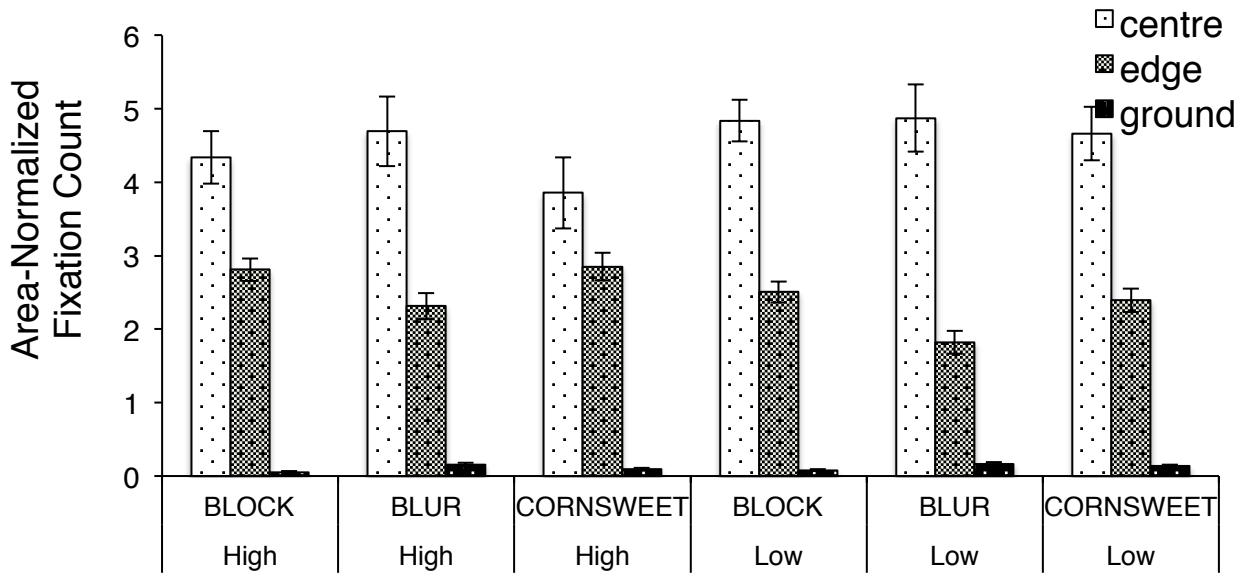


Figure 8.17. Average of all participants' data for the area-normalized proportion of fixation count for each orientation profile and orientation contrast in response to the Edge Task. Error bars represent the standard error of the mean.

Centre Task. The 2 × 3 repeated measures ANOVA (see Table 8.5 and Figure 8.18) revealed a significant main effect of orientation contrast, where Bonferroni corrected pairwise comparison shows that participants made fewer fixations to stimuli with high orientation contrast (M=5.35) compared to low orientation contrast (M= 13.46). A main effect of Profile was also observed, where fewer fixations were made to the Block profile (M=8.52), compared to the Blur profile (M=11.59) [p<0.001]. No other Bonferroni corrected pairwise comparisons were significant. These results indicate that more fixations are made to stimuli with low saliency. Finally, there was no interaction between orientation contrast and orientation profile.

Table 8.5
ANOVA summary for fixation count data that was not area normalized (top) and area normalized (bottom) for the Centre Task. For analysis with respect to the different AOIs (Region variable), area-normalized data was calculated to take into account the different sizes of the AOIs.

	df	*F*	*p*
Fixation Count			
Contrast	1,23	14.329	**0.001**
Profile	2,46	3.467	**0.040**
Contrast × Profile	2,46	0.626	0. 539
Area-Normalized Proportion of Fixation Count			
Region	1.03,23.69	852.899	**<0.001**
Contrast × Region	1.03,23.79	51.520	**<0.001**
Profile × Region	1.94,44.66	9.194	**0.001**
Contrast × Profile × Region	1.81,41.69	0.248	0.910

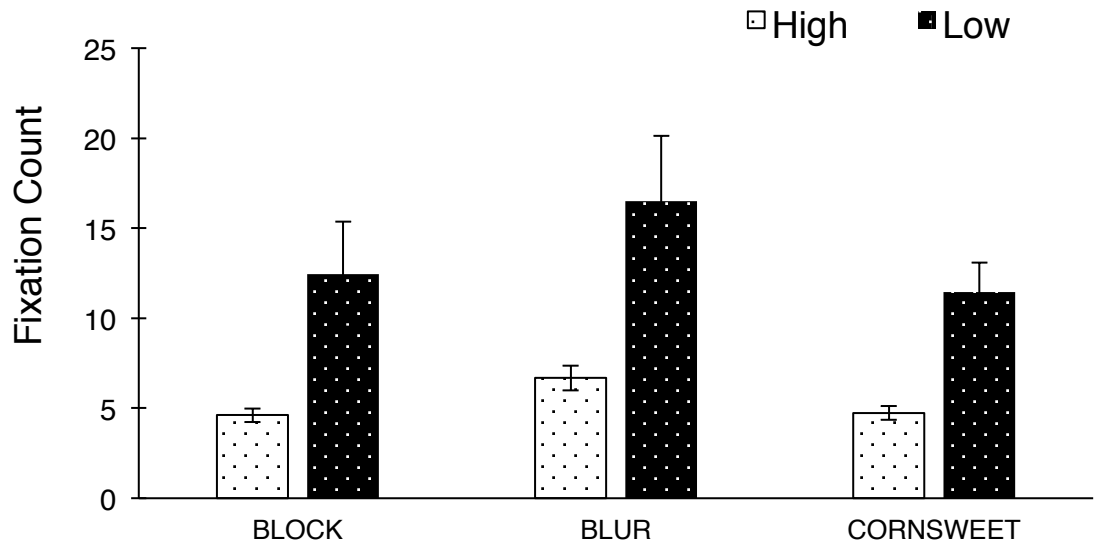


Figure 8.18. Average fixation count for the three orientation profiles across the different orientation contrast in response to the Centre Task. Error bars represent the standard error of the mean.

The following 2 × 3 × 3 repeated measures ANOVA (see Table 8.5 and Figure 8.19) was conducted to determine how fixation count varied across the different AOIs. The area-normalized proportion of fixations was used for this analysis. A significant main effect of region was found, which shows that the proportion of fixations to centre region (M=7.71)

was the highest, followed by the edge region (M=1.31), and then the ground region (M=0.09) [for all Bonferroni corrected pairwise comparisons, p<0.001].

Post hoc tests using the Bonferroni correction was used to analyse the significant contrast and region interaction. The comparisons show that as orientation contrast increases, the proportion of fixations increases to the centre region (M_{High}=8.61; M_{Low}=6.81) [p<0.001], and decreases to the edge (M_{High}=1.10; M_{Low}=1.52) [p<0.001] and ground (M_{High}=0.04; M_{Low}=0.14) [p<0.001] region. This shows that the proportion of fixations to centre region increases as the target figure becomes more detectable.

Bonferroni corrected pairwise comparisons of the profile and region interaction show that the proportion of fixations to centre region of the Blur profile (M=7.02) is lower compared to the Block (M=8.24) [p<0.001] and Cornsweet (M=7.86) [p=0.010] profile. However, the Blur profile has a higher proportion of fixations to the edge and ground region compared to the Block and Cornsweet profile. For the edge region, the proportion of fixations to the Blur profile (M=1.44) is higher than the Block (M=1.18) [p=0.015] and Cornsweet (M=1.29) [p=0.029] profile. Likewise, for the ground region, the proportion of fixations to the Blur profile (M=0.13) is higher, compared to the Block profile (M=0.06) [p=0.002] and Cornsweet profile (M=0.07) [p=0.014]. These results suggest that participants were having difficulties in localising the centre of the Blur profile, which resulted in a lower proportion of fixations to the centre, but an increase to the edge and ground.

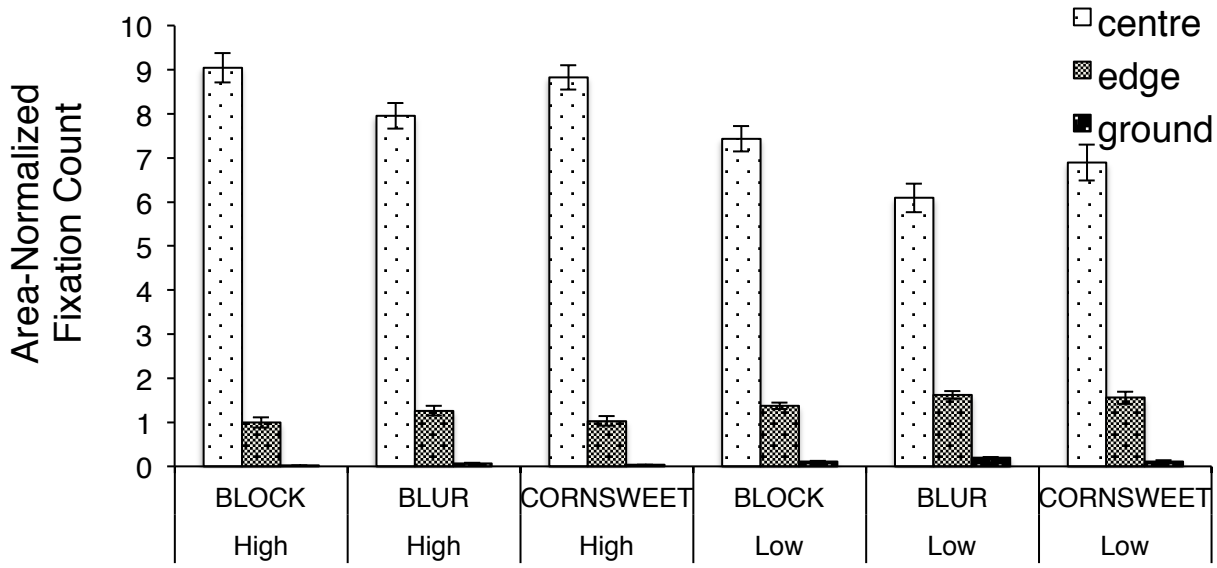


Figure 8.19. Average of all participants' data for the area-normalized proportion of fixation count for each orientation profile and orientation contrast in response to the Centre Task. Error bars represent the standard error of the mean.

8.2.3.4 Summed Fixation Duration

The summed fixation duration analyses were conducted in the same manner as the fixation count data (see Section 8.2.3.3). That is, two ANOVAs were conducted, the first on the absolute values of summed fixation duration (2-way analysis) and the second on the area normalized summed fixation duration data (3-way analysis).

For main effects and interactions in which the assumption of sphericity was violated, the Greenhouse-Geisser ($\varepsilon<0.75$) or Huynd-Feldt ($\varepsilon>0.75$) correction was applied to the degrees of freedom and the p-values.

Edge Task. The 2 × 3 repeated measures ANOVA (see Table 8.6, top and Figure 8.20) on the absolute values of summed fixation duration revealed a significant main effect of orientation contrast, in which Bonferroni corrected pairwise comparison reveal that participants made shorter total fixations to stimuli with high orientation contrast (M=1757ms) compared to low

orientation contrast (*M*=2766ms). A main effect of Profile revealed that longer total fixations were made to the Blur profile (*M*=2733ms), compared to the Block profile (*M*=1957ms) [$p<0.001$] and Cornsweet profile (*M*=2095ms) [$p=0.003$]. These results indicate that when a stimulus is more salient, the summed fixation duration is shorter.

Finally, there was a significant interaction between Contrast and Profile. Post hoc tests using Bonferroni correction show that for the high contrast stimuli, the Block profile (*M*=1472ms) was looked at for shorter durations, compared to the Cornsweet (*M*=1807ms) [$p=0.022$] and Blur profile (*M*=1992ms) [$p=0.001$]. However, for the low contrast stimuli, both the Block (*M*=2442ms) [$p<0.001$] and Cornsweet (*M*=2382ms) [$p=0.003$] profile had shorter total fixation durations compared to the Blur (*M*=3475ms). The result shows that as orientation contrast increases, the Block profile, that has information regarding orientation contrast at both figure centre and edge, has the shortest summed fixation durations.

Table 8.6
ANOVA summary for summed fixation duration data that was not area normalized (top) and area normalized (bottom) for the Edge Task. For analysis with respect to the different AOIs (Region variable), area-normalized data was calculated to take into account the different sizes of the AOIs.

	df	*F*	*p*
Summed Fixation Duration			
Contrast	1,21	80.466	**<0.001**
Profile	1.45,30.46	20.266	**<0.001**
Contrast × Profile	2,42	5.494	**0.008**
Area-Normalized Proportion of Summed Fixation Duration			
Region	1.03,23.74	73.681	**<0.001**
Contrast × Region	1.03,23.67	6.951	**0.014**
Profile × Region	1.96,45.04	8.130	**0.001**
Contrast × Profile × Region	2.10,58.37	0.200	0.830

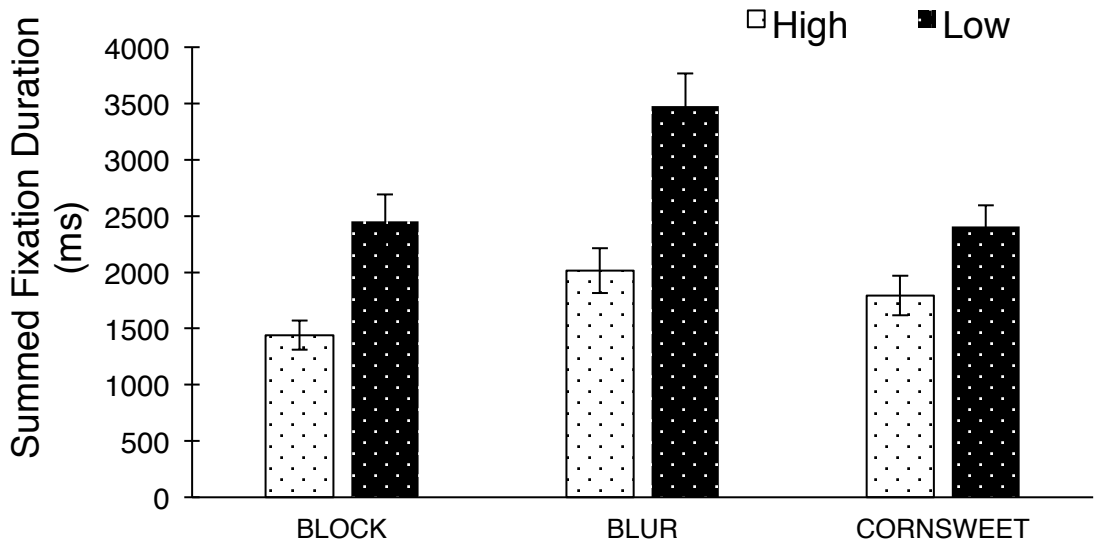


Figure 8.20. Average summed fixation duration for the three orientation profiles across the different orientation contrast in response to the Edge Task. Error bars represent the standard error of the mean.

Next, we present the 3-way analysis we conducted, which allowed us to determine how the summed fixation durations varied across the different AOIs, after taking into account the different sizes of the AOI. The ANOVA (see Table 8.6 and Figure 8.21) revealed a significant main effect of region, which shows that the proportion of summed fixation duration to centre region (M=3.93) was the highest, followed by the edge region (M=2.78), and then the ground region (M=0.09) [centre vs. edge, p=0.046, for all other Bonferroni corrected pairwise comparisons, p<0.001]. This is once again an unexpected outcome as we hypothesised that the proportion of summed fixation duration to the edge region will be greater than the centre region when performing the Edge Task.

Bonferroni corrected pairwise comparisons of the contrast and region interaction show that orientation contrast influenced the proportion of summed fixation duration to the centre (M_{High}=3.54; M_{Low}=4.43) [p=0.041] and edge region (M_{High}=3.06; M_{Low}=2.51) [p=0.001], but not the ground region (M_{High}=0.07; M_{Low}=0.10) [p=0.076]. Therefore, as orientation contrast

increases, the proportion of summed fixation duration decreases to the centre region, but increases to the edge region. Therefore, as the saliency of the stimulus increases, participants spend more time viewing the task-relevant region.

Post hoc tests using the Bonferroni correction was used to analyse the profile and region interaction. The comparisons reveal that the proportion of summed fixation duration to the centre region of the Blur profile (M=4.54) is higher compared to the Block (M=3.82) [p=0.015] and Cornsweet (M=3.45) [p=0.007] profile. For the edge region, the proportion of summed fixation duration to the Blur profile (M=2.26) is lower than the Block (M=3.05) [p<0.001] and Cornsweet (M=3.03) [p<0.001] profile, while for the ground region, the proportion of summed fixation duration to the Block profile (M=0.04) is lower than the Blur (M=0.13) [p=0.003] and Cornsweet profile (M=0.09) [p=0.020]. Thus, in comparison with the other profiles, the Blur profile shows a decrease in the proportion of summed fixation duration to the edge, but an increase to the centre and ground. This could be due to an increased difficulty in localising the edge of the Blur figure.

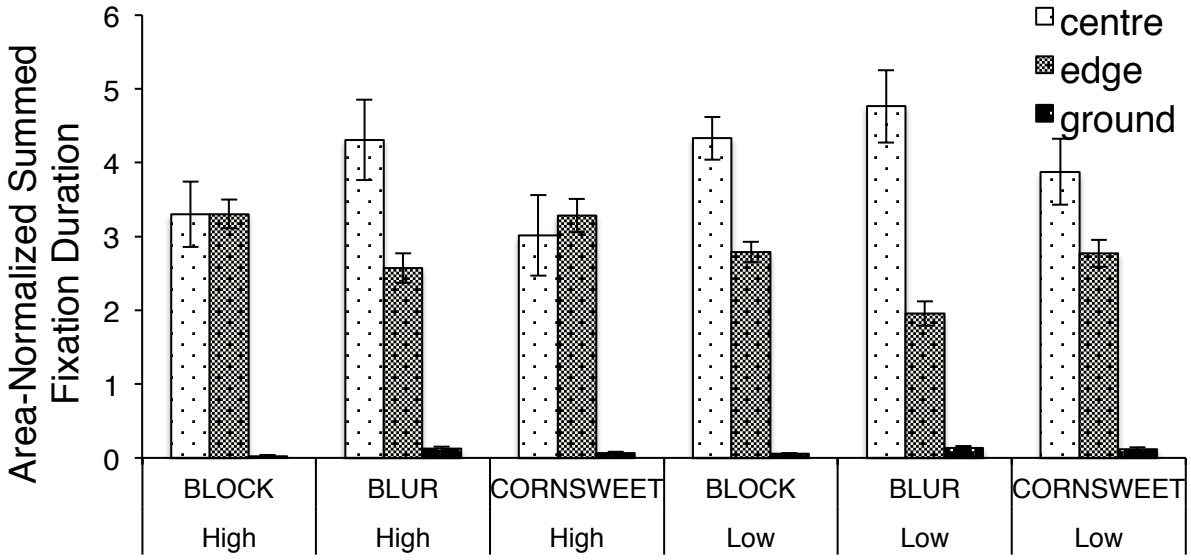


Figure 8.21. Average of all participants' data for the area-normalized proportion of summed fixation duration for each orientation profile and orientation contrast in response to the Edge Task. Error bars represent the standard error of the mean.

Centre Task. The 2 × 3 repeated measures ANOVA (see Table 8.7 and Figure 8.22) revealed a significant main effect of orientation contrast, where Bonferroni corrected pairwise comparison show that participants made shorter total fixations to stimuli with high orientation contrast (M=1632ms) compared to low orientation contrast (M=3327ms). A main effect of Profile revealed that longer total fixations were made to the Blur profile (M=2933ms), compared to the Block profile (M=2315ms) [p<0.001]. No other Bonferroni corrected post hoc tests were significant. These findings suggest that shorter summed fixation durations are made to targets that are more salient. Finally, there was no interaction between orientation contrast and orientation profile.

Table 8.7
ANOVA summary for summed fixation duration data that was not area normalized (top) and area normalized (bottom) for the Centre Task. For analysis with respect to the different AOIs (Region variable), area-normalized data was calculated to take into account the different sizes of the AOIs.

	df	*F*	*p*
Summed Fixation Duration			
Contrast	1,23	13.056	**0.001**
Profile	2,46	3.725	**0.032**
Contrast × Profile	2,46	1.213	0.307
Area-Normalized Proportion of Summed Fixation Duration			
Region	1.01,23.26	560.784	**<0.001**
Contrast × Region	1.04,23.76	79.648	**<0.001**
Profile × Region	1.92,44.18	9.293	**0.001**
Contrast × Profile × Region	1.79,41.27	1.510	0.206

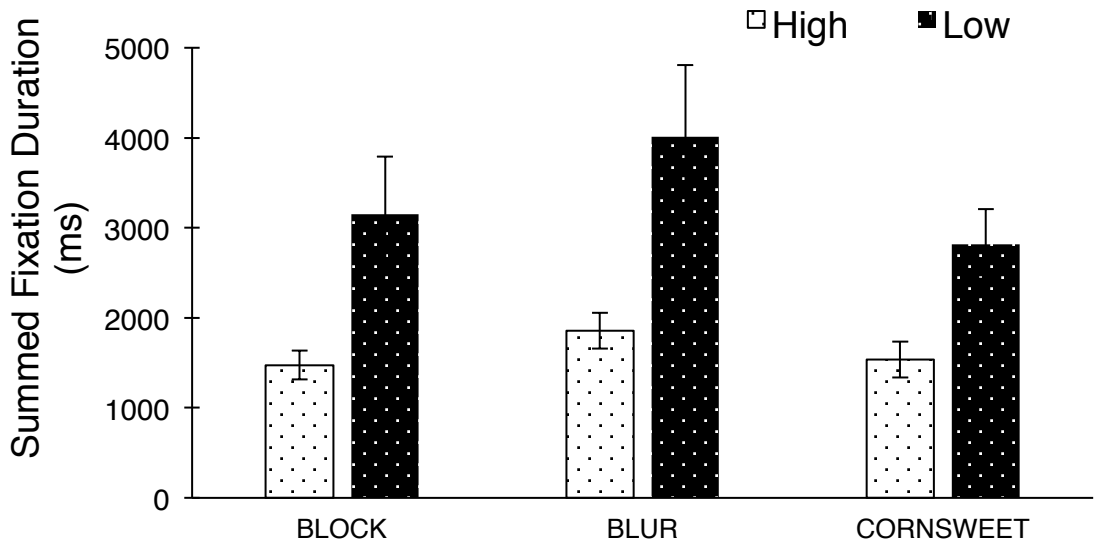


Figure 8.22. Average summed fixation duration for the three orientation profiles across the different orientation contrast in response to the Centre Task. Error bars represent the standard error of the mean.

The following 2 × 3 × 3 repeated measures ANOVA (see Table 8.7 and Figure 8.23) was conducted to determine how the summed fixation duration varied across the different AOIs. The area-normalized proportion of summed fixation duration was used for this analysis. A significant main effect of region shows that the proportion of summed fixation duration to centre region (M=8.57) was the highest, followed by the edge region (M=1.03), and then the ground region (M=0.07) [for all pairwise comparisons using Bonferroni correction, p<0.001].

Post hoc tests using the Bonferroni correction was used to analyse the significant contrast and region interaction. The comparisons show that as orientation contrast increases, the proportion of summed fixation duration increases to the centre region (M_{High}=9.58; M_{Low}=7.56) [p<0.001], and decreases to the edge (M_{High}=0.77; M_{Low}=1.29) [p<0.001] and ground (M_{High}=0.02; M_{Low}=0.11) [p<0.001] region. This shows that participants spend a longer proportion of time viewing the centre region when orientation contrast increases.

Bonferroni corrected pairwise comparisons of the profile and region interaction reveal that the proportion of summed fixation duration to centre region of the Blur profile (M=7.88) is lower compared to the Block (M=9.12) [p=0.001] and Cornsweet (M=8.72) [p=0.008] profile. However, the Blur profile has a higher proportion of summed fixation duration to the edge and ground region compared to the Block and Cornsweet profile. For the edge region, the proportion of summed fixation duration to the Blur profile (M=1.21) is higher than the Block (M=0.88) [p=0.016] and Cornsweet (M=0.99) [p=0.030] profile. Likewise, for the ground region, the proportion of summed fixation duration to the Blur profile (M=0.10) is higher, compared to the Block profile (M=0.04) [p<0.001] and Cornsweet profile (M=0.06) [p=0.015]. It is possible that participants are having difficulties with localising the centre

region of the Blur profile, which has resulted in a lower proportion of summed fixation duration to the centre region.

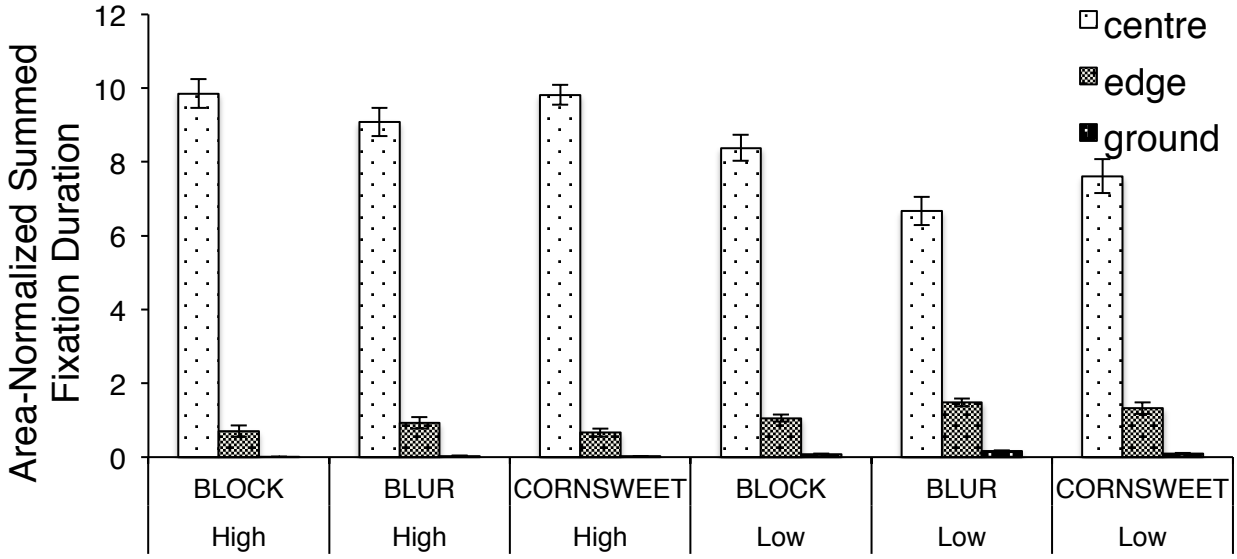


Figure 8.23. Average of all participants' data for the area-normalized proportion of summed fixation duration for each orientation profile and orientation contrast in response to the Centre Task. Error bars represent the standard error of the mean.

8.3 Discussion

Both the experiments presented in this chapter were conducted with the aim of investigating 1) where in a texture stimulus do people look, 2) specifically, if eye movements to texture stimuli are affected by the different orientation profiles, and 3) if eye movements to texture stimuli are affected by orientation contrast.

To address the first aim, we analysed the landing position of the first fixation. Initially, this was done using the RoFF data, which is the proportion of fixations to each region AOI. However, further analysis of Experiment 7 using the positional coordinates of first fixation revealed that participants' primary fixation was most likely being directed to the centre of the entire texture grid, with an undershoot of approximately 15%. Therefore, in this experiment,

participants did not demonstrate an ability to localise the target figure within the first fixation. This was unexpected, as studies that used texture stimuli have shown that humans (e.g. Deubel et al., 1988) and even animals (e.g. Supèr, Spekreijse, & Lamme, 2001) are able to localise target figures within the first fixation. However, it is possible that the outcome was due to the design of the experiment. The task in our study was to localise the target figure and identify its centre/edge **if** the figure was present. Additionally, the experiment used low orientation contrast stimuli that had target figures that were not easily detectable. This could have resulted in participants adopting a strategy where they carefully scan the entire texture. This experimental design is different compared to the aforementioned studies that used texture stimuli with figures that pop-out from the background. As such, in the following chapter, we will present a series of experiments in which the methodology was specifically altered to address the first aim mentioned here.

In the studies reported in this chapter, two other eye tracking measures were used: fixation count and summed fixation duration. We found a higher number of fixations and longer summed fixation durations when orientation contrast decreased. Moreover, the Block profile constantly had fewer fixations and shorter summed fixation durations. Therefore, as target saliency increased, the number and duration of fixations decreased. These results are possibly a consequence of participants spending more time examining a target figure when it is not very detectable. This effect of increased fixation count to low saliency targets is supported by Van Humbeeck et al., (2013). The authors asked participants to search for a contour target, and found that the number of fixations and total viewing time increased when the target was less salient.

Both the fixation count and summed fixation duration measures were also used because regions of importance or interest will have a greater number of fixations (Holmqvist et al.,

2011), and longer summed fixation durations (Janik et al., 1978). As such, we expected that participants would fixate more often and for longer times on the regions of a texture figure that produce the most information regarding orientation contrast between figure and ground. Specifically, we predicted a greater proportion of fixations to the centre region of the Blur profile, and a greater proportion of fixations to the edge region of the Cornsweet profile. For the Block profile, we expected an equal proportion of fixations to the centre and edge region, owing to both regions containing information regarding orientation contrast. Additionally, we expected the tasks in Experiment 7 to influence the frequency and summed duration of fixations. That is, the task demand was expected to produce a greater proportion of fixations to the task-relevant feature.

However, we did not find the predicted pattern of results. In fact, the area normalized results show that the centre region is looked at more often and for longer durations, irrespective of the orientation profiles or task. However, while the area-normalized proportion of fixation count and summed fixation duration to the edge region is lower than the centre region, the area-normalized data still shows that fixations to the edge region is specifically targeted. That is, even though the edge is viewed less often and for shorter times compared to the centre, the fixations to the edge region are occurring at a rate higher than that expected by chance. We are able to conclude this because the data used were area normalized i.e. proportion responses were divided by the proportional size of the region. The ground region however constantly has proportion values lower than one, which is what one would expect since participants were able to localise the figure. Besides that, the centre region had a significantly higher proportion of fixation count and summed fixation duration compared to the edge region for both the Edge Task and Centre Task. However, the proportion of fixations to the edge region did increase when participants performed the Edge Task. For the area-normalized proportion of fixation count to the edge region, it was 1.31 for the Centre Task, and 2.45 for the Edge

Task. For the area-normalized proportion of summed fixation duration to the edge region, it was 1.03 for the Centre Task, and 2.78 for the Edge Task. Therefore, even though the centre region was viewed the most often and for longer times, the inclusion of a task-demand to locate the edge increases the proportion of fixation count and summed fixation duration to the edge region.

Furthermore, the orientation contrast affects the proportion of fixation count and summed fixation duration to the different regions. Specifically, we find that as orientation contrast increases, both measures show a decrease in response to the centre region and an increase to the edge region for the Edge Task. In contrast, there is an increase in response to the centre region and a decrease to the edge region for the Centre Task. This suggests that as orientation contrast increases and the target becomes more detectable, the proportion of fixations to the task-relevant feature increases. Likewise, the orientation profile affects the proportion of fixation count and summed fixation duration to the different regions. In particular, the Blur profile shows a decrease in the proportion of fixation count and summed fixation duration to the task-relevant feature compared to the Block and Cornsweet profile. It is possible that participants are finding it more difficult to localise the Blur figure compared to the other profiles.

In Experiment 7, we investigated the time it takes to initiate a saccade to the primary fixation. According to some studies (e.g. Supèr, Spekreijse, & Lamme, 2003; Supèr & Lamme, 2007), faster saccades are elicited due to stronger figure-ground signals. In the case of our stimuli, the low orientation contrast stimuli would have a weaker figure-ground signal, which should have resulted in longer saccade latencies. However, our results are not in line with this as orientation contrast did not have an effect on saccade latencies. A possible explanation for this is that the first fixation was not being directed to the figure, but to a point just left of the

centre of the entire texture grid. As a result, the saccade initiation was not in response to the target figure, which is why the saccade latency shows no effect of orientation contrast, profile, or task.

Besides the eye tracking measures, task performance was also measured in Experiment 7 in terms of accuracy and time to respond. Because our objective in this experiment was to ascertain whether or not participants could localize the boundary of the figure, the response we used to measure this was a mouse-click. Initially, we intended for participants to locate the centre/edge of the figure, and subsequently **maintain** fixation at that region until stimulus offset. However, when we were piloting the experiment, live-view of the participants' eye movements showed that they were having difficulties in maintaining fixation on the centre/edge of the figure. These participants were nevertheless able to localise the figure when the experimenter asked. As such, we chose to use the mouse-click as the measure for this experiment. Live-view (and also replay) of the participants' eye movements when they performed this task showed that participants located the centre/edge of the figure before moving the cursor to make a response. Moreover, weighted gaze samples of all the recordings were between 92% and 98% (mean of 96%), which indicates that participants were not looking away from the monitor to look at the mouse when making the response.

Accuracy was measured as participants' ability to correctly identify the presence or absence of a target figure, as well as the ability to correctly localise the edge/centre of the figure. Participants demonstrated improved performance with highly salient targets, which are stimuli with high orientation contrast or the Block profile. In fact, even as orientation contrast decreases, performance for the Block profile is always better than the Blur and Cornsweet profile. This could be evidence that both edge and surface information of a target figure is used to localise the target. In line with this, the time to mouse-click shows that participants

respond faster to stimuli with high orientation contrast. Additionally, performance is best for the Block profile, suggesting once again that both edge and surface information produces the highest levels of saliency for target localisation.

In conclusion, the key findings of the experiments presented are: 1) irrespective of orientation profiles or the task to locate centre/edge of the figure, participants looked at the centre region more frequently and for longer durations compared to the edge region when compared on an area-scaled basis, 2) the number of fixations and summed fixation duration decrease with increased figure saliency, 3) increased saliency of targets result in a higher proportion of fixations to the task-relevant feature, 4) compared to the Block and Cornsweet profiles, the Blur profile showed a decrease in proportion of fixations to the task-relevant feature, 5) task performance improves when the target figure is more salient, 6) the Block profile with orientation contrast at both figure centre and edge is most salient, and 7) the position of first fixation is to a point just left of the centre of the texture grid, and this position appears to be unrelated to the target figure for this type of experimental setup.

Chapter 9

Investigating The Effect of Target Position and Orientation Profile on Eye Movements to Texture Stimuli

In this chapter, we present three experiments (Experiments 8, 9, & 10) that were conducted with the general aim of investigating how the different edge orientation profiles affect eye movements, and also where in a texture stimulus participants would look. The results from Experiment 7 (Chapter 8) showed that the region to which participants made their first fixation had very little to do with the stimulus profiles (Block, Blur, and Cornsweet) or orientation contrast (high and low contrast). However, it is possible that this outcome could be attributed to the experimental design itself. The results of that experiment suggest that participants did not localise the figure region within the first saccade. This is in contrast to other studies (e.g. Deubel et al., 1988; Supèr, Spekreijse, & Lamme, 2001) that have found that it is possible to make saccadic eye movements to targets defined by a change in orientation from the background. We therefore had inconclusive results about the effect on eye movements that the various orientation profiles have.

Thus, in Experiment 8, we once again used texture stimuli with the three different orientation profiles, but in this experiment, all the stimuli used were easily detectable i.e. high orientation contrast only. We varied the eccentricities of the target and got participants to make an immediate eye movement to the target once it appeared. In Experiment 9, we decreased the size and increased the eccentricity of the target in order to investigate whether saccade amplitudes will be increased sufficiently to ensure that the primary fixation lands on the target figure. A follow-up question we had was whether saccades were planned to the centre

of the target, or the centre of gravity (Melcher & Kowler, 1999). To investigate this, we changed the shape of the target figure in Experiment 10 to a triangle-shaped target instead of the rectangle-shaped figure used thus far.

9.1 Experiment 8: Eye Movements to High Orientation Contrast Stimuli with Different Orientation Profiles at Varying Eccentricities

A particular limitation of the previous studies (Chapter 8) was that half of the stimuli used were of low orientation contrast. Due to individual differences, some of the participants may well have been unable to localise the figure. That is, participants whose thresholds were far below the contrast used would have been able to easily detect the figure. On the other hand, participants whose thresholds were around or above the contrast used would not have been able to detect the figure easily. This is because for participants whose thresholds were around the contrast used, the figure would have been just slightly noticeable, and some of the eye movements made would be the participant searching for the figure. Whereas, for participants whose thresholds were above the contrast used, figure detection would have been immediate. This variation of participants' thresholds could have introduced noise to the data.

The current study used stimuli with the same profiles as the previous studies (Block, Blur, and Cornsweet), but in this case, the target (figure) was always clearly visible i.e. considerably above threshold at 90° orientation contrast. To address the issue in the previous experiment of participants always making a primary fixation to a location just left of the centre of the texture grid, the experimental design was changed in a number of ways. First, the fixation point was in the centre of the screen, with the target appearing on either the left or right side of fixation (see Deubel & Frank, 1991). Second, we limited the stimulus duration to 1 second (instead of 5 seconds, and terminating upon response for the Pilot study

and Experiment 7 respectively). By doing this, we forced the participants to make immediate eye movements to the clearly visible targets instead of giving them the opportunity to scan for the target. Finally, instead of randomly placing the target within the texture grid, we systematically varied the position of the targets, whereby it could appear at three different eccentricities on either side of fixation. The purpose of having multiple possible locations is so the participants could not predict the target position. The participants therefore had to localise the figure region **before** making a saccade to the target.

Based on previous literature, we expect that saccade latencies will be greater when the targets are further from the fovea (e.g. Land & Tatler, 2009; Gilchrist, 2011), and the stimulus is less salient (e.g. Sumner, 2011; Supèr, Spekreijse, & Lamme , 2003). However, the more novel findings we expected was for the landing position of the primary fixation for orientation defined targets. If eye movements are driven by the most salient features in a scene (e.g. Itti & Koch, 2000; Parkhurst et al., 2003), we should expect to see great variations in landing positions for the different edge profiles. Specifically, the Blur and Cornsweet profiles should have primary fixations closer to the centre and edge of the target figures respectively (as that is where orientation contrast information is present), while the Block profile will have primary fixations throughout the entire figure. However, if saccades are driven by a representation of a whole figure (post integration of elements into segmented region) as suggested by Melcher and Kowler (1999) or perhaps a spatial pooling mechanism (e.g. Deubel et al., 1988; Ploner et al., 2004; Findlay 1982), there should be no difference between the landing position of the different profiles.

9.1.1 Methods

9.1.1.1 Participants

10 participants (5 males) between the ages of 19 and 24 (M_{age} = 21.6) were recruited from the University of Nottingham Malaysia. All participants had normal or corrected-to-normal vision.

9.1.1.2 Apparatus & Display

A Tobii T60 Eye-Tracker running Tobii Studio 3.2 was used to present the stimuli and record participants' eye movements. Participants sat 60cm away from monitor, which produced a resolution of 1.52 arcminute per pixel. Each participant performed a nine-point calibration before the experimental task.

The stimuli presented were 1024×768 pixels, which was of a texture array that had a visual angle of 25.9°×19.5°. The figure patch spanned 5.7°×4.6° (height × width), while the fixation point was 0.4° wide.

9.1.1.3 Stimuli

The stimuli were pre-generated using PsychoPy 1.83.04 (Pierce, 2007) on an Apple Mac Mini. The overall texture array comprised 50 × 68 line elements, with a figure patch that was 15×12 (height × width) elements. The figure patch could appear at six possible locations – at eccentricities of 9.1°, 6.8°, and 4.6° to the left or right of fixation (Figure 9.1). The three different edge profiles: Block, Blur, and Cornsweet were used in this study. The parameters of the stimuli constructed are the same as described in Section 3.3. The orientation contrast between figure and ground was held constant at 90°, while the orientation of the background elements (baseline orientation) was either 0° (vertical), 45°, 90°, or 135°.

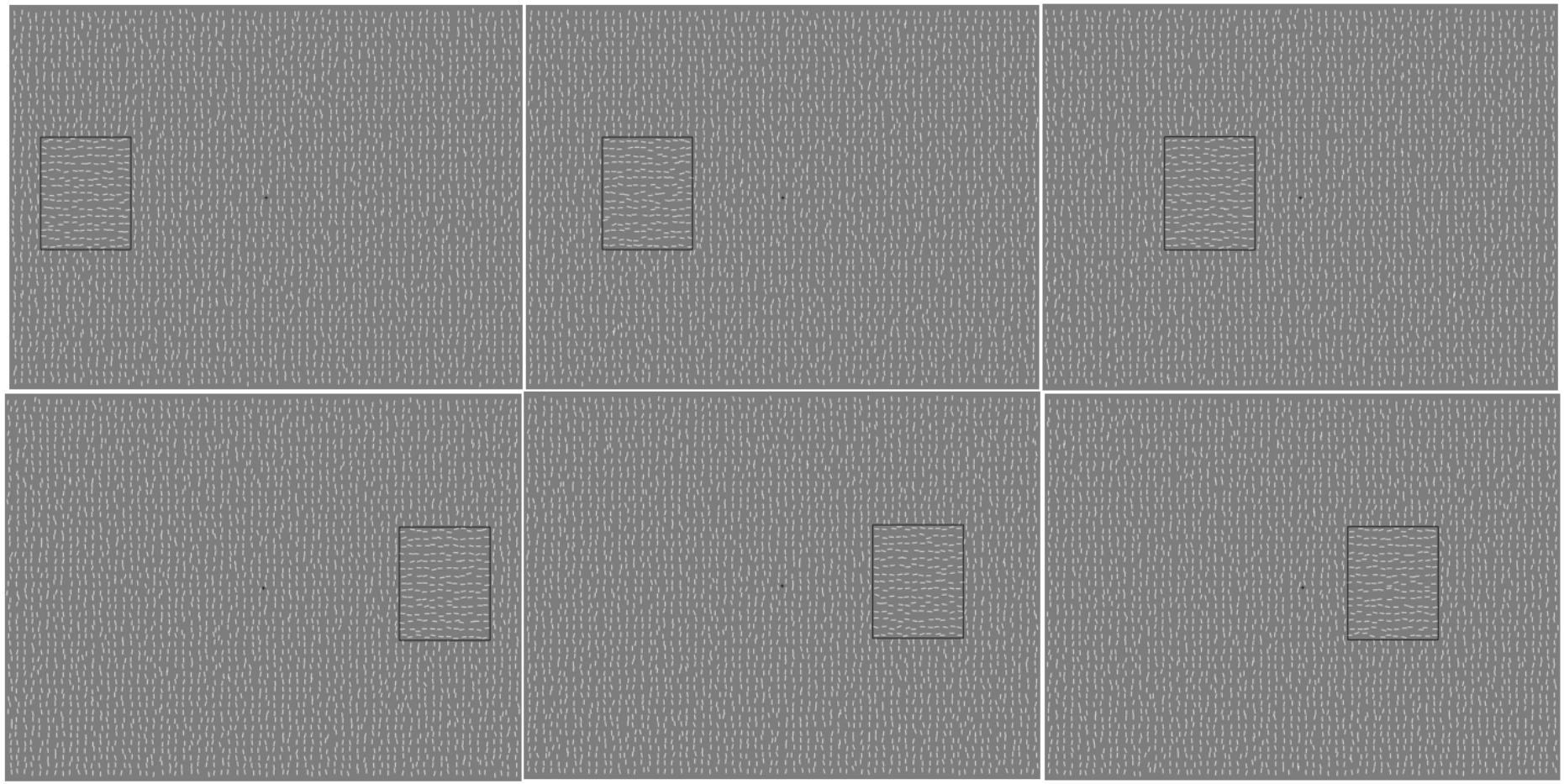


Figure 9.1. The stimuli here show examples of the position of the target figure to the left (top row) and right (bottom row) of fixation. Three eccentricities were used: 9.1° (left column), 6.8° (middle column), and 4.5° (right column). The black solid line represents the border of the target figure. All stimuli shown here are of the Block profile with 90° orientation contrast (baseline orientation contrast at 0°).

9.1.1.4 Procedure

Before the experimental session, the experimenter showed the participants examples of the stimuli. These examples were different to those that were used during the experimental session. The purpose of this was because the participants were not familiar with texture stimuli, and the verbal description of 'target figure' might be ambiguous to them. These example stimuli were shown to the participants on the eye tracker monitor, and the experimenter explicitly indicated where the target figure was located. Participants were allowed to view the example stimuli for as long as they wanted, but they were informed that in the practice and experimental runs, the texture stimuli would only remain on-screen for 1000ms.

Participants were told to make an eye movement to the target. Specifically, they were instructed: *"once the textured image appears, all you have to do is move your eyes as fast as possible to the figure"* i.e. they were not told to look at any specific point within the target figure. In the previous studies, participants were not told to maintain fixation at the target. However in this study, they were required to maintain fixation at the target until the stimulus disappeared (1000ms after stimulus onset). The instructions given to the participants were *"once you have moved your eyes to the figure, maintain your gaze on the figure until the textured image goes off"*. The target was always detectable by the participant as the orientation contrast was high (90 degrees). Participants performed 1 practice block before completing 5 experimental blocks (72 trials per block). The sequence of an experimental trial begun with a fixation cross appearing at the centre of the screen for 1000ms, followed by the stimulus for another 1000ms. After stimulus offset, the fixation cross reappeared, and the next trial was initiated.

9.1.2 Results

Before the results were analysed, trials in which participants failed to make a primary fixation to the figure centre or edge (as defined by Figure 3.10 in Chapter 3) were removed. This included trials in which participants made a first fixation to: 1) the opposite direction of target location, 2) the correct direction, but not reaching the target edge/centre, and 3) regions above/below the fixation. These excluded trials amounted to 18.1% of the total trials. The various eye tracking measures used in this study (Section 3.6.1) and the method for analysing the results (Section 3.6.2) were described earlier in Chapter 3.

9.1.2.1 Position of First Fixation

In Experiment 7 (Chapter 8), we found that participants were making the first fixation to a point just slightly left to the middle of the entire texture grid, and this point did not appear to be much influenced by the location of the embedded texture patch. However, the experimental setup may well have produced this outcome. Hence, this study was designed whereby the figure position was varied systematically to test if the figure position and profile affected the landing position of the first fixation.

The position to where participants made the first fixation was analysed in terms of percentage amplitude (%), which was calculated using the formula:

$$Saccade\ Amplitude = 100\,.\frac{Distance\ of\ primary\ fixation\ from\ fixation\ point}{Distance\ of\ target\ centre\ from\ fixation\ point}$$

As such, 100% amplitude is when the first fixation was to the exact figure centre (for all three eccentricities). However, the percentage amplitude of the figure border differs based on the eccentricity. For the target at 9.1° eccentricity, the border closer to fixation is 75% amplitude and the border further from fixation is 125%. For the target at 6.8° eccentricity, the border

closer to fixation is 66.7% amplitude and the border further from fixation is 133.3%. Finally, the target at 4.6° eccentricity, the border closer to fixation is 50% amplitude and the border further from fixation is 150%. These borders are represented as the dashed lines on the graphs in Figure 9.3.

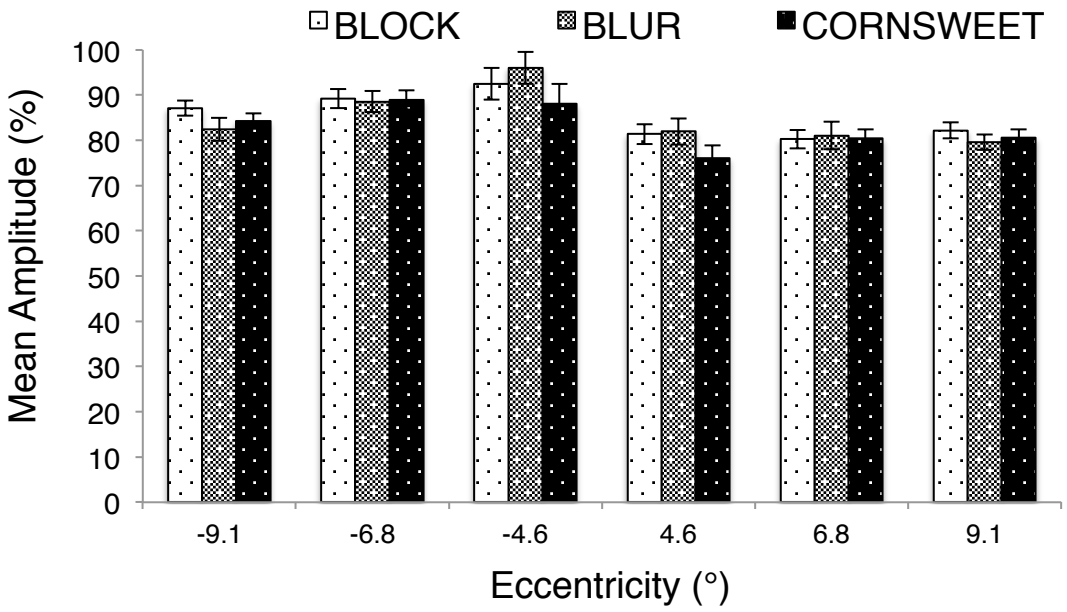
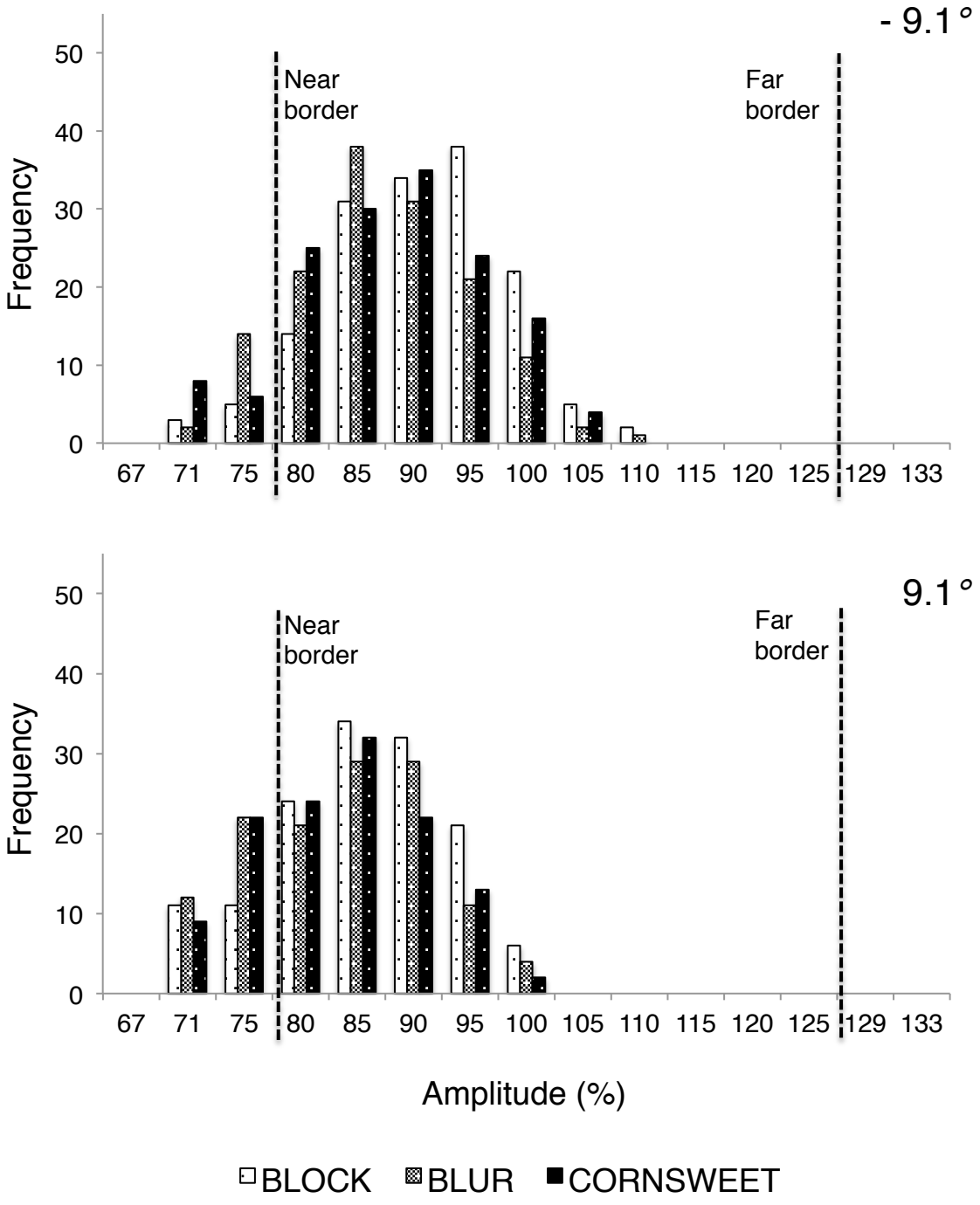


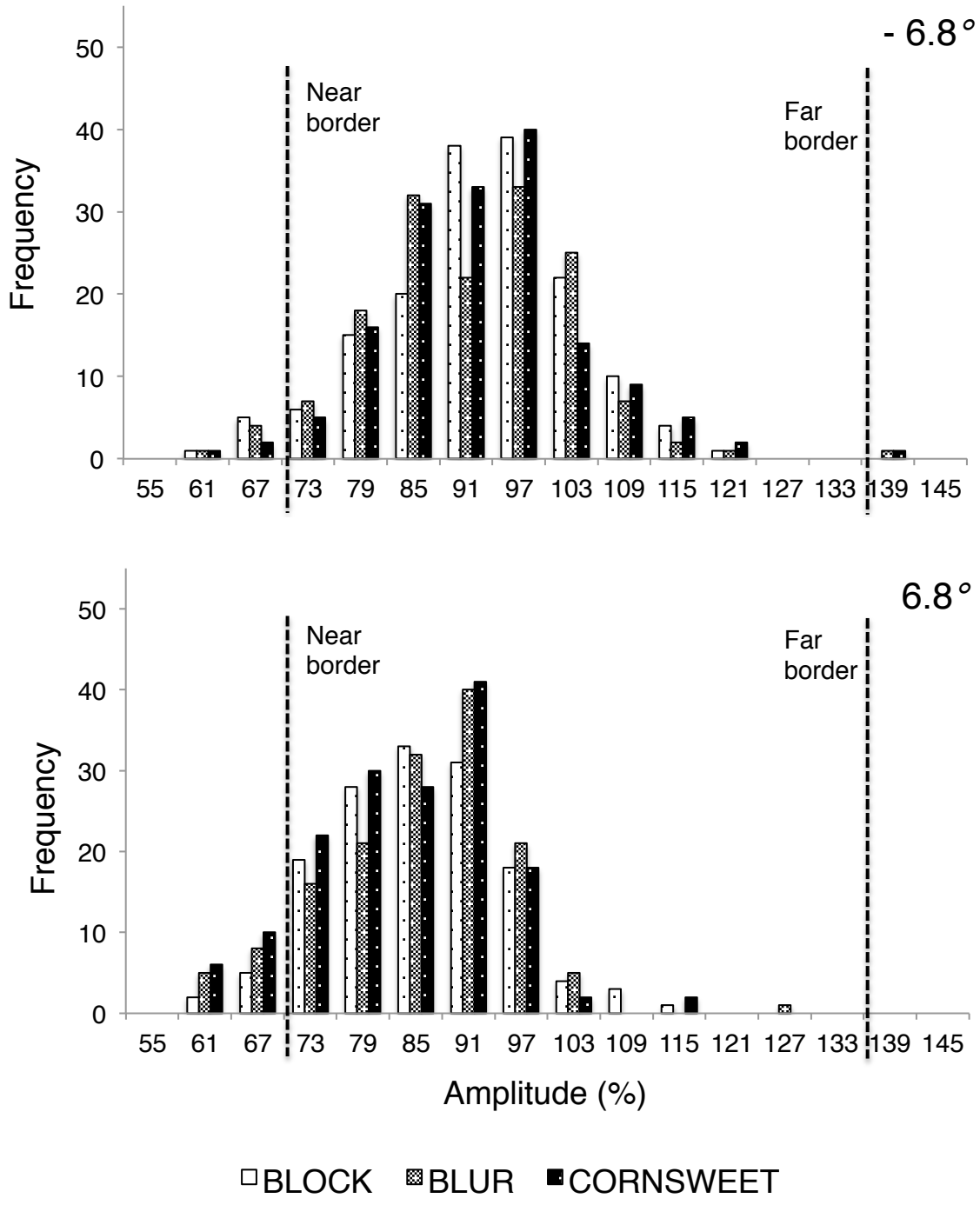
Figure 9.2. Average amplitudes of first fixation in percentage relative to the centre of figure (100%). Negative (-) eccentricities represent targets to the left of fixation. Error bars represent the standard error of the mean.

Figure 9.2 shows the mean amplitude for all the profiles at all six positions. While the mean amplitude shows very little change between the profiles and position, this is only in terms of amplitude data. In fact, for absolute pixel units, the different positions have different values, with larger values for targets further away from fixation. This is expected, as participants would have to make a larger saccade to a target that is further away. But overall, participants are making a saccade to the target, with an approximate 15% undershoot.

In order to visually inspect the pattern in the data, the percentage amplitude of all trials for all the participants were plotted as frequency histograms for each profile at each location (see

Figure 9.3). Visual inspection of the frequency histogram shows that the profiles do not have an effect on the distribution of the amplitude data i.e. the profiles are not affecting the landing position of the first fixation. A few other observations can be made from the frequency histograms. First, the amplitude of the landing position of the first fixation seems to be on average lower than 100% (figure centre). Second, the amplitude distribution shows that the landing positions are typically made around 80-90%. Third, very rarely did first fixations land at amplitudes greater than 100%, especially for targets further away from fixation. The implication of the aforementioned points is that when a primary fixation is made, the landing position is typically on a point just slightly short of the figure centre, but not falling on the figure edge either. Finally, when the target is to the right of the fixation point, the amplitude distribution also seems lower compared to when the target is to the left. The percentage amplitude data was also analysed using a 4-way ANOVA (See Figure 9.2), the outcome of which is shown in Table 9.1. The independent variables of the ANOVA are: Left-Right – figure could appear on the left or right side of fixation; Position – the figure could be at one of three locations (4.6°, 6.8°, or 9.1°); Profile – Block, Blur, and Cornsweet); and Region – Centre or Edge. For main effects and interactions in which the assumption of sphericity was violated, the Greenhouse-Geisser ($\varepsilon<0.75$) or Huynd-Feldt ($\varepsilon>0.75$) correction was applied to the degrees of freedom and the p-values.


- 9.1°
Near border
Far border
Frequency
50
40
30
20
10
0
67 71 75 80 85 90 95 100 105 110 115 120 125 129 133
9.1°
Near border
Far border
Frequency
Amplitude (%)
BLOCK
BLUR
CORNSWEET


- 6.8°
Near border
Far border
Frequency
50
40
30
20
10
0
55 61 67 73 79 85 91 97 103 109 115 121 127 133 139 145
6.8°
Near border
Far border
Frequency
50
40
30
20
10
0
55 61 67 73 79 85 91 97 103 109 115 121 127 133 139 145
Amplitude (%)
BLOCK
BLUR
CORNSWEET

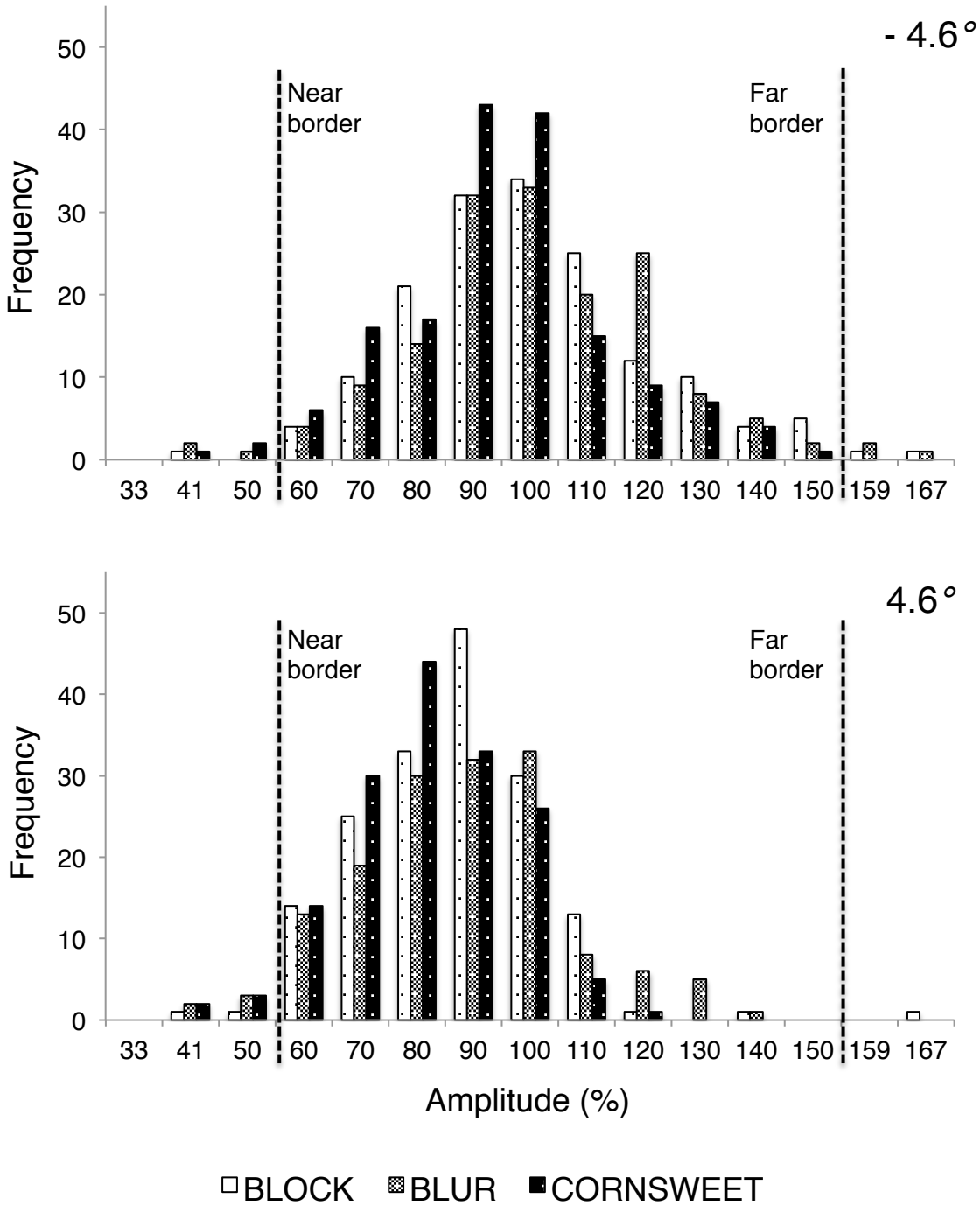


Figure 9.3. Saccade amplitude data (in percentage relative to figure centre) plotted as frequency histograms, with pooled data from all 10 participants. The extremes of the x-axis scale on the graphs correspond to the edge region (2 columns away from the border). The eccentricity and direction of the target figure is indicated on the top right corner of the graph, and the dashed lines on the graph represent the border of the figure. The intervals on the x-axis represent the upper limit of the bin array e.g. 80 includes amplitude up to 80%.

Table 9.1
Summary of ANOVA results on Position of First Fixation (saccade amplitude in percentage relative to figure centre).

	df	*F*	*p*
Left-Right	1,9	14.824	**0.004**
Position	1.173,10.558	1.626	0.224
Profile	2,18	4.337	**0.029**
Left-Right × Position	2,18	7.506	**0.004**
Left-Right × Profile	2,18	0.037	0.964
Position × Profile	2.169,19.519	5.092	**0.015**
Left-Right × Position × Profile	4,36	1.075	0.383

There was no main effect of Position, which shows that the position of the target did **not** influence the percentage saccade amplitude relative to the centre of the target figure. A significant main effect of Left-Right was found, which showed that mean amplitudes were lower for targets to the right (M=80.40%) compared to the left (M=88.56%) [where p=0.004 for Bonferroni corrected pairwise comparison]. These results show the same pattern as the frequency histograms above, where the overall amplitude is lower for targets to the right of fixation point, and that on average the amplitudes are around 80-90%. Furthermore, a main effect of Profile found that amplitudes were lower for the Cornsweet profile (M=83.06%) than the Block profile (M=85.45%) [p=0.010]. A point to note about this is that even though the amplitudes to the Cornsweet profile is lower on average, it is still not close to the figure defined edge (in fact it is closer to the centre). No other Bonferroni corrected post hoc tests were significant.

Bonferroni corrected pairwise comparisons of the interaction between Left-Right and Position show that position does not influence the amplitude to targets to the right (where p>0.05), but for targets to the left, amplitude is lower for targets at 9.1° eccentricity (M=84.60%) compared to targets at 6.8° eccentricity (M=88.90%) [p=0.042]. Finally, an interaction between Position and Profile shows that amplitude is affected by the profiles at

4.6° target eccentricity, in which the amplitude was significantly lower to the Cornsweet profile than the Block [p=0.008] and Blur [p=0.022] profiles (M_{Block}=86.95%; M_{Blur}=88.99%; $M_{Cornsweet}$=82.1%). Amplitude is not affected by the profiles at the 6.8° target eccentricity (M_{Block}=84.76%; M_{Blur}=84.81%; $M_{Cornsweet}$=84.67%) or the 9.1° target eccentricity (M_{Block}=84.63%; M_{Blur}=81.00%; $M_{Cornsweet}$=82.43%) [p>0.05 for all Bonferroni corrected post hoc comparisons].

9.1.2.2 Precision of Landing Position

The precision of the landing position was measured by calculating the standard deviation of the saccade amplitudes. This was initially done per participant and per condition. Subsequently, the standard deviations were analysed with a 2 (left-right: target to left or right of fixation) × 3 (position: 9.1°, 6.8°, and 4.6°) × 3 (profile: Block, Blur, and Cornsweet) repeated measures ANOVA (see Figure 9.4 and Table 9.2). For main effects and interactions in which the assumption of sphericity was violated, the Greenhouse-Geisser (ε<0.75) or Huynd-Feldt (ε>0.75) correction was applied to the degrees of freedom and the p-values.

Table 9.2
Summary of ANOVA results on Precision (Standard Deviation) data.

	df	*F*	*p*
Left-Right	1,9	3.630	0.089
Position	2,18	26.887	**<0.001**
Profile	1.193,10.741	1.181	0.313
Left-Right × Position	2,18	0.132	0.877
Left-Right × Profile	1.216,10.946	1.088	0.335
Position × Profile	4,36	1.208	0.324
Left-Right × Position × Profile	1.978,17.802	1.563	0.237

Table 9.2 above shows the outcome of the ANOVA for the precision data. A significant main effect of Position was found, in which Bonferroni corrected pairwise comparisons reveal that

the landing positions of the first fixation for the targets at 4.6° eccentricity was more scattered (M=16.5%) compared to targets at 9.1° (M=10.1%) [p<0.001] and 6.8° eccentricity (M=8.0%) [p=0.002]. Thus, the reliability of the saccadic landing position was better when the target was further away from fixation. No other comparisons were significant.

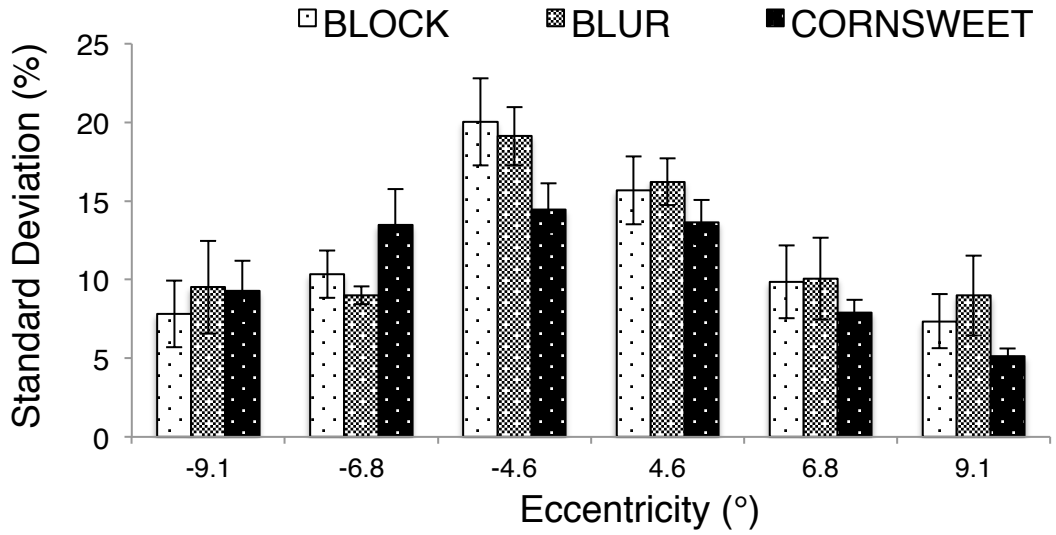


Figure 9.4. Average standard deviation of the first fixation as a function of target profile and eccentricity. Negative (-) eccentricities represent targets to the left of fixation. Error bars represent the standard error of the mean.

9.1.2.3 Saccade Latency

The saccade latency is a measure of time duration for a participant to initiate a saccade. The time measurement starts when the stimulus is first displayed, and stops when the participant initiates an eye movement. A 2 (left-right: target to left or right of fixation) × 3 (position: 9.1°, 6.8°, and 4.6°) × 3 (profile: Block, Blur, and Cornsweet) repeated measures ANOVA (see Figure 9.5 and Table 9.3) was carried out. For main effects and interactions in which the assumption of sphericity was violated, the Greenhouse-Geisser (ε<0.75) or Huynd-Feldt (ε>0.75) correction was applied to the degrees of freedom and the p-values.

Table 9.3
Summary of ANOVA results on Saccade Latency data.

	df	*F*	*p*
Left-Right	1,9	0.183	0.679
Position	2,18	9.709	**0.001**
Profile	2,18	37.920	**<0.001**
Left-Right × Position	2,18	0.882	0.431
Left-Right × Profile	2,18	1.096	0.356
Position × Profile	2.058,18.520	1.379	0.261
Left-Right × Position × Profile	4,36	0.805	0.530

Table 9.3 above shows the outcome of the ANOVA. Bonferroni corrected pairwise comparisons of the main effect of Position show that saccade latencies for targets at 9.1° eccentricity (M=304ms) was longer compared to targets at 4.6° (M=293ms) [p=0.025] and 6.8° eccentricity (M=293ms) [p=0.013]. A main effect of Profile indicates that participants have longer saccade latencies for the Blur profile (M=311ms) compared to the Block (M=291ms) and Cornsweet profile (M=289ms) [for all Bonferroni corrected pairwise comparisons, p<0.001].

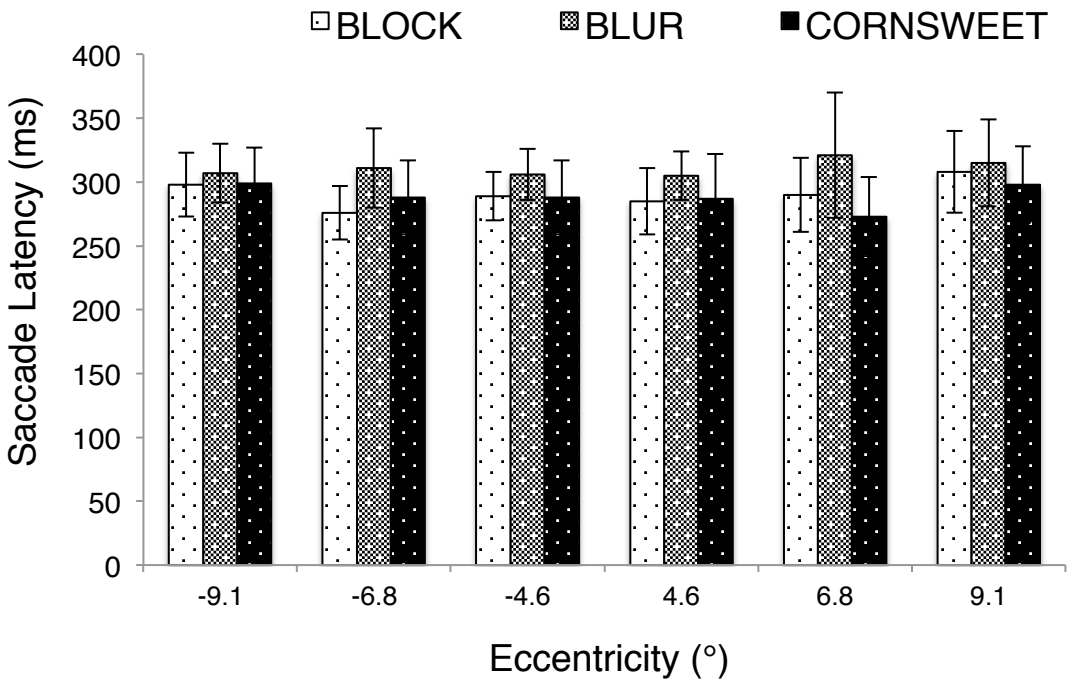


Figure 9.5. Average saccade latencies as a function of target profile and eccentricity. Negative (-) eccentricities represent targets to the left of fixation. Error bars represent the standard error of the mean.

9.1.2.4 Fixation Count

To investigate the effects of target position and target profile on the number of fixations on the stimuli, two ANOVAs were carried out. First, a 2 × 3 × 3 (left-right × position × profile) ANOVA, and subsequently a 2 × 3 × 3 × 2 (left-right × position × profile × region: centre, and edge) ANOVA. The 3-way ANOVA was carried out using the absolute values of the fixation count data, while the 4-way ANOVA was carried out using the area-normalized fixation count data (see Section 3.6.2 for area-normalizing). For the 4-way ANOVA, the main effects and interaction of left-right, position, and profile variables were not reported as they were explored in the 3-way analysis.

For main effects and interactions in which the assumption of sphericity was violated, the Greenhouse-Geisser ($\varepsilon<0.75$) or Huynd-Feldt ($\varepsilon>0.75$) correction was applied to the degrees of freedom and the p-values.

Table 9.4
ANOVA summary for fixation count data that was not area normalized (top) and area normalized (bottom). For analysis with respect to the different AOIs (Region variable), area-normalized data was calculated to take into account the different sizes of the AOIs.

	df	*F*	*p*
Fixation Count			
Left-Right	1,9	0.580	0.466
Position	2,18	1.638	0.222
Profile	2,18	1.093	0.374
Left-Right × Position	2,18	1.564	0.236
Left-Right × Profile	2,18	1.876	0.182
Position × Profile	4,36	2.957	0.066
Left-Right × Position × Profile	4,36	0.771	0.551
Area-Normalized Proportion of Fixation Count			
Region	1,9	226.137	**<0.001**
Left-Right × Region	1,9	1.231	0.296
Position × Region	2,18	1.481	0.254
Profile × Region	2,18	5.604	**0.013**
Left-Right × Position × Region	2,18	1.947	0.172
Left-Right × Profile × Region	2,18	1.881	0.181
Position × Profile × Region	4,36	2.052	0.108
Left-Right × Position × Profile × Region	2.40,21.58	1.614	0.220

The 3-way ANOVA (see Table 9.4 top and Figure 9.6) showed that none of the main effects or interactions were significant. Thus, not the position of the target, nor orientation profile influenced the number of fixations on the stimuli.

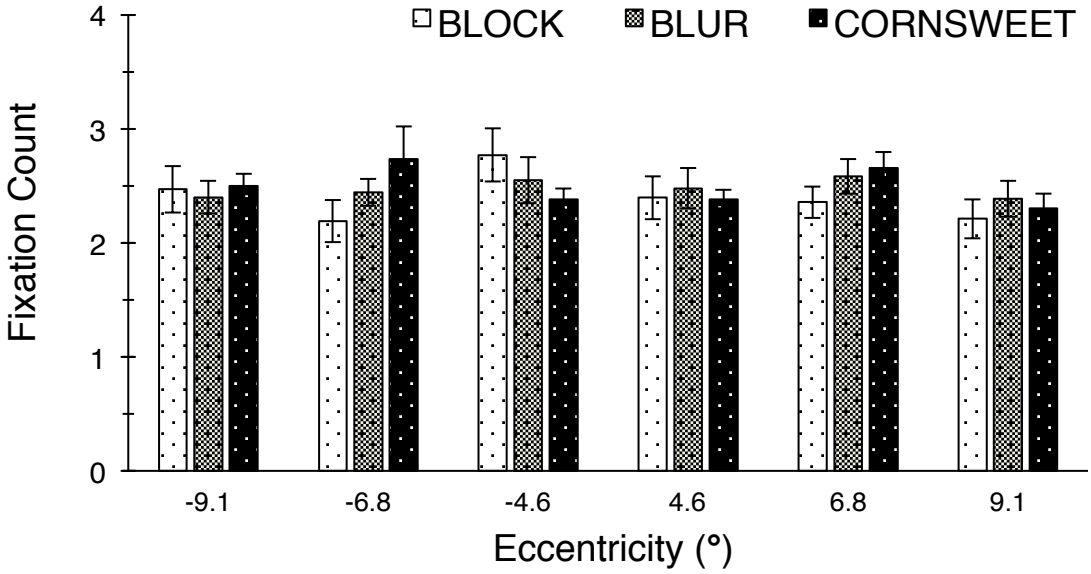


Figure 9.6. Average fixation count (absolute value) for the three profiles across the different target positions. Negative (-) eccentricities represent targets to the left of fixation. Error bars represent the standard error of the mean.

The 4-way ANOVA (see Table 9.4 bottom and Figure 9.7) showed a significant main effect of Region. The area-normalized proportion of fixations to the centre region centre region (M=21.78) was greater than the edge region (M=6.87) [where p<0.001 for the pairwise comparison using Bonferroni correction]. This indicates that fixations are directed more frequently to the centre region of the target figure compared to the edge region. However, as we discussed in Section 3.6.2, any value greater than 1 indicates that the proportion of fixations to that region is above what one would expect by chance alone. Therefore, the result suggests that fixations directed to the edge region are specifically targeted (i.e. not occurring due to chance alone), even though it is at a disproportionately smaller extent compared to the centre region.

Bonferroni corrected pairwise comparisons of the interaction between Region and Profile found that for the response to the centre region, the proportion fixations to the Block profile was higher compared to the Blur profile (M_{Block}=23.08; M_{Blur}=20.87) [p=0.001]. However,

for response to the edge region, the opposite is seen, where the proportion fixations to the Blur profile was higher compared to the Block profile (M_{Block}=6.34; M_{Blur}=7.24) [p=0.001]. The Blur profile was created to have very little information of orientation change at the edge of the figure – yet we find that the proportion of fixations made to the edge region is higher for the Blur profile than the Block. However, the proportion of fixations to the centre region is still greater than the edge region for the Blur profile (which is what we expect).

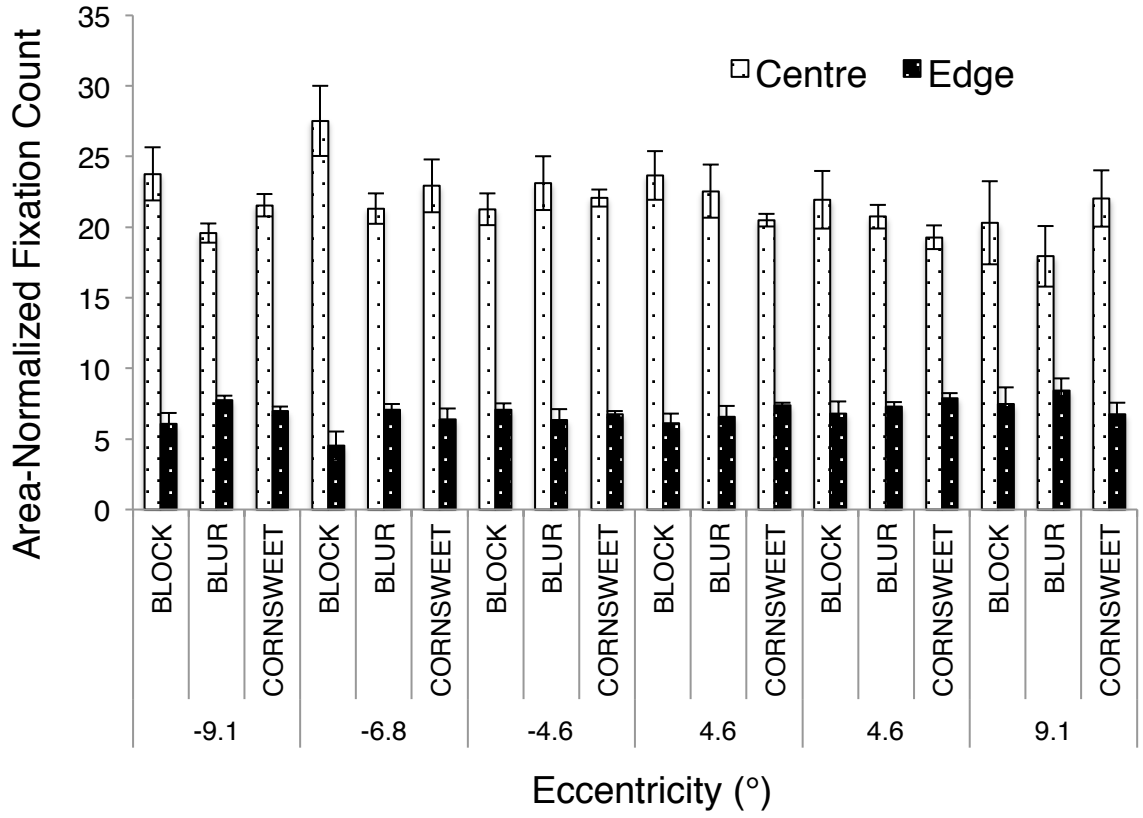


Figure 9.7. Mean area-normalized proportion of fixation count (proportion fixation count to AOI divided by proportion AOI size), where Centre and Edge are the pre-defined AOIs. Negative (-) eccentricities represent targets to the left of fixation. Error bars represent the standard error of the mean.

9.1.2.5 Summed Fixation Duration

To investigate the effects of target position and target profile on the summed fixation duration to the stimuli, two ANOVAs were carried out. First, a 2 × 3 × 3 (left-right × position

× profile) ANOVA, and subsequently a 2 × 3 × 3 × 2 (left-right × position × profile × region: centre, and edge) ANOVA. The 3-way ANOVA used the absolute values of the summed fixation duration data, while the 4-way ANOVA used the area-normalized summed fixation duration data (see Section 3.6.2 for area-normalizing). As with the fixation count analysis, we did not explore the main effects and interaction of the left-right, position, and profile variables of the 4-way ANOVA as they were explored in the 3-way analysis.

For main effects and interactions in which the assumption of sphericity was violated, the Greenhouse-Geisser ($\varepsilon<0.75$) or Huynd-Feldt ($\varepsilon>0.75$) correction was applied to the degrees of freedom and the p-values.

Table 9.5
ANOVA summary for the summed fixation duration data that was not area normalized (top) and area normalized (bottom). For analysis with respect to the different AOIs (Region Variable), area-normalized data was calculated to take into account the different sizes of the AOIs.

	df	*F*	*p*
Summed Fixation Duration			
Left-Right	1,9	0.179	0.682
Position	1.246,11.210	0.119	0.790
Profile	2,18	0.258	0.776
Left-Right × Position	2,18	0.163	0.851
Left-Right × Profile	2,18	0.408	0.671
Position × Profile	1.896,17.066	0.463	0.627
Left-Right × Position × Profile	4,36	1.130	0.358
Area-Normalized Proportion of Summed Fixation Duration			
Region	1,9	116.814	**<0.001**
Left-Right × Region	1,9	2.713	0.134
Position × Region	2,18	8.713	**0.002**
Profile × Region	1.09,9.77	1.874	0.203
Left-Right × Position × Region	2,18	0.764	0.480
Left-Right × Profile × Region	2,18	0.259	0.774
Position × Profile × Region	1.68,15.12	0.363	0.666
Left-Right × Position × Profile × Region	4,36	0.500	0.736

The 3-way ANOVA (see Table 9.5 top and Figure 9.8) showed that none of the main effects or interactions were significant. Thus, not the position of the target, nor orientation profile influenced the summed fixation durations to the stimuli.

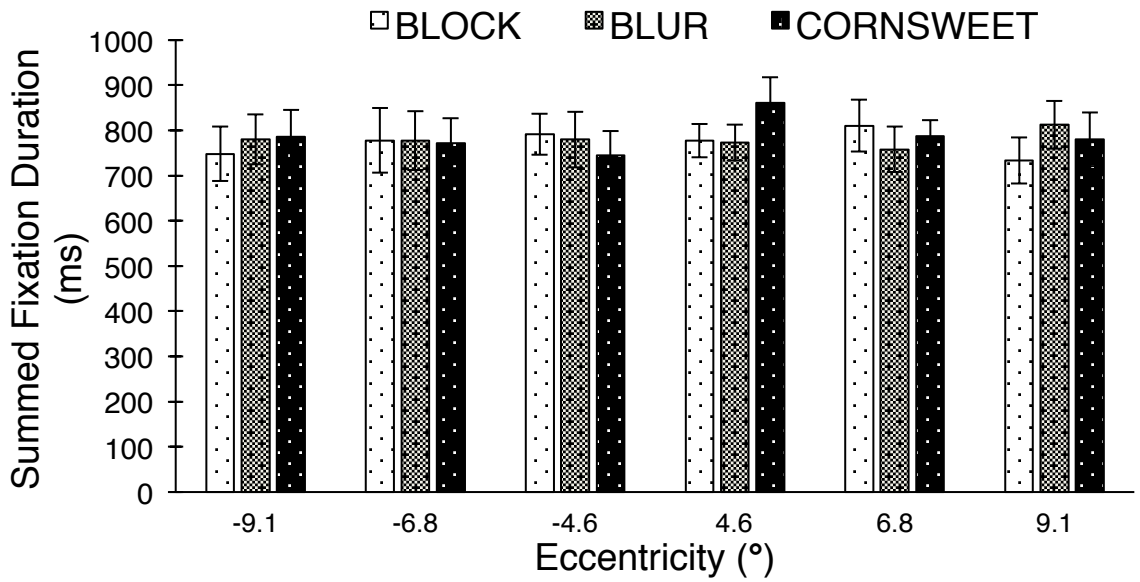


Figure 9.8. Average summed fixation duration (absolute value) for the three profiles across the different target positions. Negative (-) eccentricities represent targets to the left of fixation. Error bars represent the standard error of the mean.

The 4-way ANOVA (see Table 9.5 bottom and Figure 9.9) found a significant main effect of Region, whereby the area-normalized proportion of summed fixation durations to the centre region (M=23.68) was greater than the edge region (M=6.09) [where $p<0.001$ for the pairwise comparison using Bonferroni correction]. Once again, even though the scaled proportion of fixation time to the centre region far exceeds the edge region, the edge region is still looked at for a proportion of time that is more than what one would expect from chance alone. This is the same pattern of results we found with the fixation count data.

For the position and region interaction, Bonferroni corrected post hoc tests show that for the centre region, the proportion of summed fixation duration was the lowest for targets at 9.1°

eccentricity compared to targets at 4.6° [p=0.035] and 6.8° [p=0.017] eccentricities ($M_{4.6°}$=25.12; $M_{6.8°}$= 24.46; $M_{9.1°}$=21.46). However, for the edge region, we find an opposite effect, where the proportion response was highest for the 9.1° targets compared to the 4.6° [p=0.035] and 6.8° [p=0.017] targets ($M_{4.6°}$=5.50; $M_{6.8°}$= 5.77; $M_{9.1°}$=7.00).

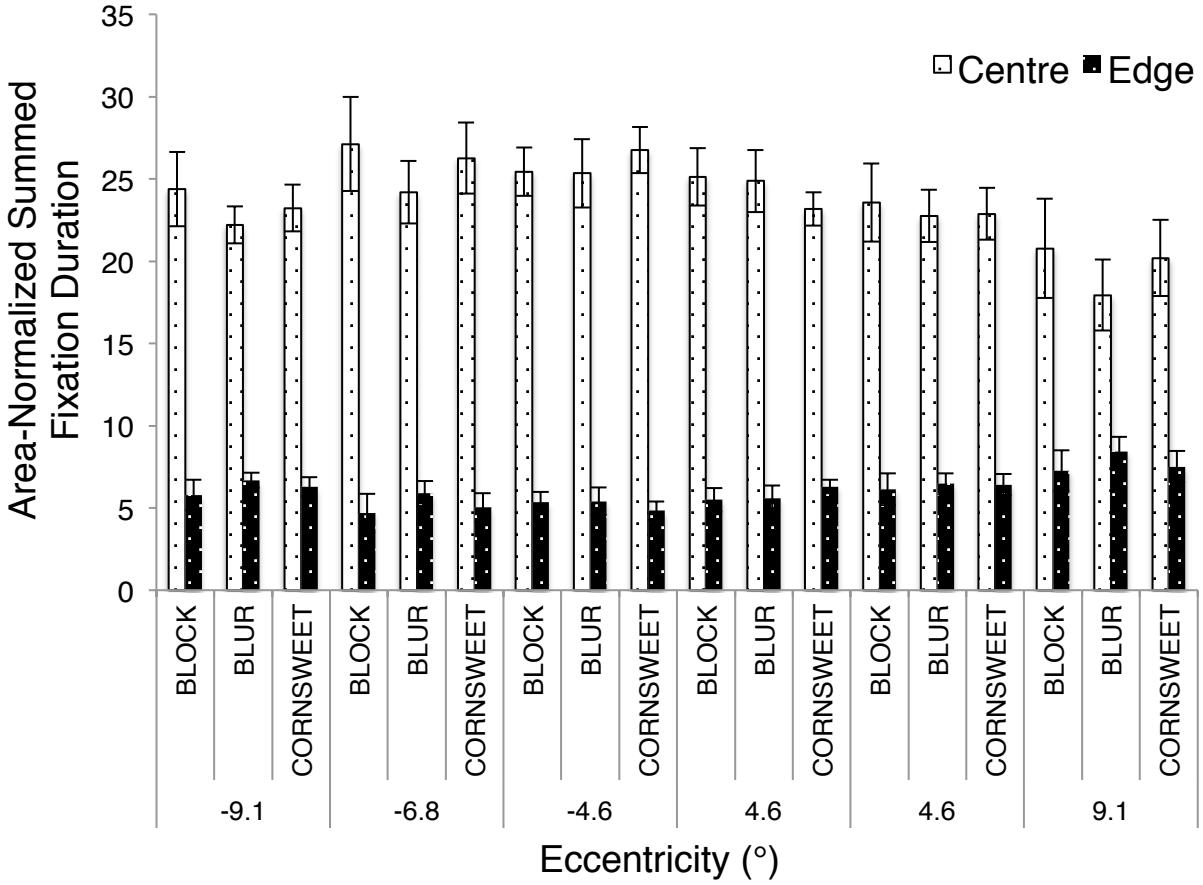


Figure 9.9. Mean area-normalized proportion of summed fixation duration (proportion summed fixation duration to AOI divided by proportion AOI size), where Centre and Edge are the pre-defined AOIs. Negative (-) eccentricities represent targets to the left of fixation. Error bars represent the standard error of the mean.

9.2 Experiment 9: Eye Movements to Targets at High Eccentricity

In Experiment 8, we found that participants generally make a primary fixation to the centre of a target figure, with an undershoot of about 15%. Typically, this resulted in the primary

fixation landing within the border of the target figure. However, we were interested in whether the amplitude of the primary fixation will always be approximately 85%, even if it resulted in the primary fixation landing on the background instead of the target figure. To test this, we made the target figure smaller and increased the target eccentricity. This resulted in the 85% amplitude landing position falling outside the figure border (i.e. on the background of the figure-ground texture stimuli) for the high eccentricity target. If participants' amplitudes are maintained at 85% for the high eccentricity target, it would imply that saccades are planned by first determining the landing position (the centre of the target), followed by a saccade with an undershoot of approximately 15% that always stays the same regardless of the stimulus. However, if amplitudes for the high eccentricity targets increased (i.e. > 85%), it would imply that saccadic movement could be changed to compensate for the additional distance required.

9.2.1 Methods

9.2.1.1 Participants

10 participants (4 males) between the ages of 21 and 28 (M_{age} = 22.9) were recruited from the University of Nottingham Malaysia. All participants had normal or corrected-to-normal vision.

9.2.1.2 Apparatus & Display

The apparatus setup and display were the same as Experiment 8.

The stimuli presented were 1024×768 pixels, which was of a texture array that had a visual angle of 25.9°×19.5°. The figure patch spanned 5.7°×1.5° (height × width), with the width size of this figure being $\frac{1}{3}$ the size of the figure width of Experiment 8. The position of the

centre of the figure patch varied at eccentricities of 6.8° (same as Experiment 8) or 10.6° to the left or right of fixation. The fixation point was 0.4° wide.

9.2.1.3 Stimuli

The stimuli were pre-generated using PsychoPy 1.83.04 (Pierce, 2007) on an Apple Mac Mini. The overall texture array comprised 50 × 68 line elements, with a figure patch that was 15×4 (height × width) elements. The figure patch could appear at four possible locations – at eccentricities of 6.8° and 10.6° to the left or right of fixation (Figure 9.10). The orientation of the line elements were constructed to form the Block profile (see Section 3.3), with orientation contrast of 90° between figure and ground elements. The orientation of the ground elements (baseline orientation) were either 0° (vertical), 45°, 90°, and 135°.

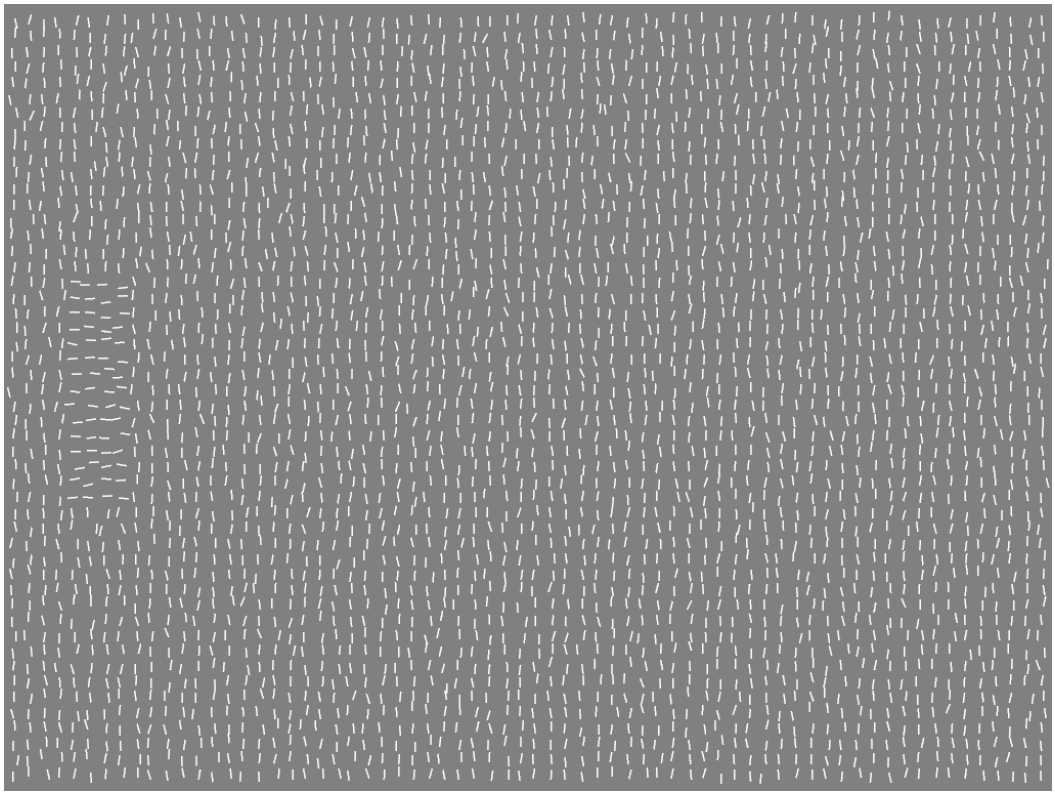


Figure 9.10. Stimuli used in the experiment. The target figure is of a Block profile with 90° orientation contrast at 10.6° eccentricity to the left of fixation. The targets in the experiment appeared at eccentricities of 6.8 and 10.6° to the left or right of a central fixation point.

9.2.1.4 Procedure

The instructions to the participants and the sequence of an experimental trial were the same as in Experiment 8 (Section 9.1.1.4). However, in the present study, participants performed 2 experimental blocks (40 trials per block).

9.2.2 Results

In the previous experiment (Experiment 8), if participants did not make a primary fixation to the centre or edge (two columns on either side of the border), the trials were removed from the analysis. However, in this current study, the aim was to investigate whether the saccade landing position falls on the target figure by having saccades with amplitudes > 85%, or will saccades have amplitudes of approximately 85% with the landing position of the saccade falling outside the border of the target figure. In consideration of the fact that the saccade landing position in this experiment may possibly fall **beyond** the border and the two columns outside of the border, we accepted trials in which the landing position was within the border, and also four columns outside the border.

9.2.2.1 Position of First Fixation

Percentage amplitude (%) was calculated (see Section 9.1.2.1) for both target eccentricities of 6.8° and 10.6°. In pixel units, the distance between the central fixation and target centre is 270 and 420. For the target at 6.8° eccentricity, the amplitudes of the figure border closest and furthest from fixation are 89% and 111%. For the target at 10.6° eccentricity, that amplitude is 93% and 107% for the near and far border.

Figures 9.11 and 9.12 shows the mean amplitudes of the participants' first fixation, as well as the frequency histograms for the various conditions. From the graphs, we observe a slight

effect of target eccentricity on amplitude. A 2 (left-right: target to left or right of fixation) × 2 (position: 6.8° and 10.6° eccentricity) ANOVA on the mean amplitude data corroborates this, as we found a main effect of target position [$F(1,9)=49.583$, where $p<0.001$] in which Bonferroni corrected pairwise comparison shows that targets at 6.8° eccentricity had a mean amplitude that was higher ($M=98.71\%$) compared to targets at 10.6° eccentricity ($M=94.43\%$). There was no main effect of Left-Right [$F(1,9)=0.047$, where $p=0.833$] or an interaction effect [$F(1,9)= 0.047$, where $p=0.927$].

Compared to the amplitudes we found in Experiment 8, which were approximately 85%, the amplitudes of first fixation we find in this experiment is higher at approximately 95%. Therefore, participants were able to localise the figure with a greater degree of accuracy in this experiment. One could claim that the increased amplitudes to the high eccentricity targets were due to saccades having greater amplitudes to ensure that the primary fixation fell within the figure border. However, it does not explain why the amplitudes were **also** increased (98%) for the nearer targets. In fact, for the 6.8° eccentricity targets, the only difference between the current experiment and Experiment 8 with the comparable Block profile is the width size of the figure. Despite this, the average amplitude to targets at 6.8° eccentricity was 84.8% for Experiment 8, and 98.7% for the present experiment. This suggests that the higher saccade amplitudes were driven by the smaller size of the figure, and not the goal of increasing the amplitudes to make certain that the location of first fixation fell within the figure border.

Nevertheless, Figure 9.12 can provide some insight into the landing position of the primary fixations. Inspection of the figure shows that a greater number of fixations landed outside the border of the figure (dashed lines on the graph) for the high eccentricity targets. For the 6.8° eccentricity target, 8% of trials fell outside the figure-defined border. For the 10.6°

eccentricity target, 29% of trials fell outside the figure-defined border. This shows that while participants were able to make a saccade that landed on the figure for most of the trials, when the target eccentricity was high, a higher portion of trials fell outside the figure border.

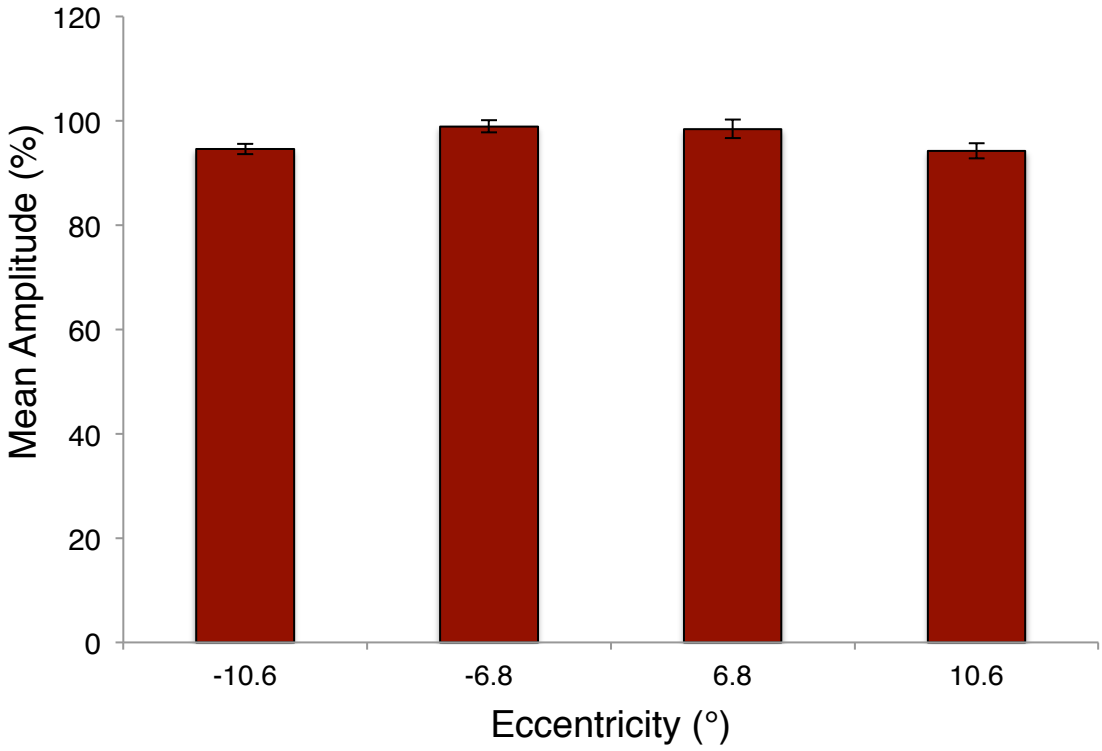


Figure 9.11. Average amplitudes of primary fixation in percentage relative to centre of figure (100%). Negative (-) eccentricities represent targets to the left of fixation. Error bars represent the standard error of the mean.

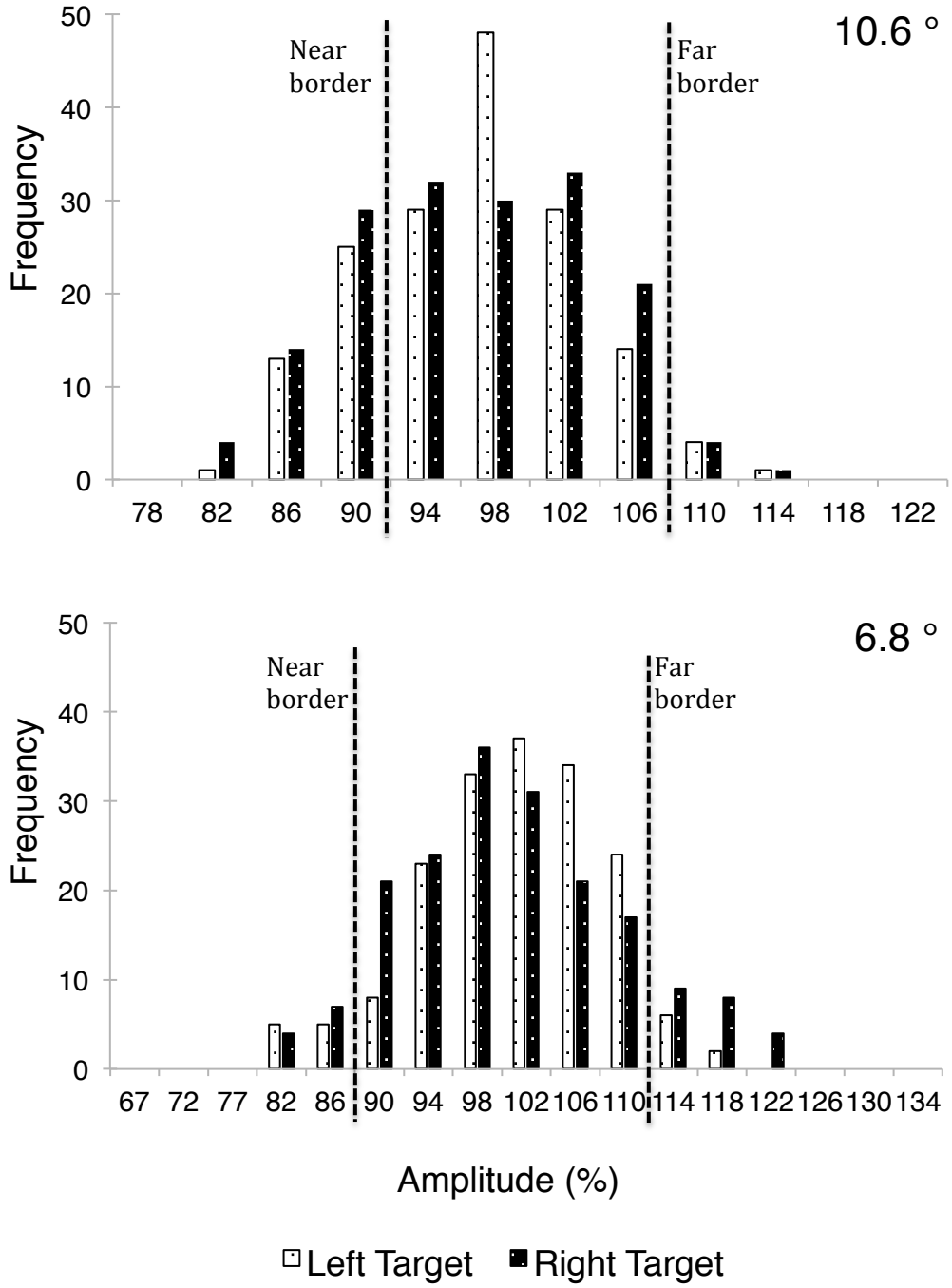


Figure 9.12. Saccade amplitude data (in percentage relative to figure centre) plotted as frequency histograms, with pooled data from all 10 participants. The extremes of the x-axis scale on the graphs correspond to the edge region (4 columns away from the border). The eccentricity of the target figure is indicated on the top right corner of the graph, and the dashed lines on the graph represent the border of the figure. The intervals on the x-axis represent the upper limit of the bin array e.g. 98 includes amplitude up to 98%.

9.2.2.2 Precision of Landing Position

The precision of the landing position was analysed with a 2 (left-right: target to left or right of fixation) × 2 (position: 10.6°, and 6.8°) repeated measures ANOVA (see Figure 9.13). A main effect of target position was found [$F(1,9)=8.230$, where $p=0.019$], in which pairwise comparison using Bonferroni correction reveals that targets at 6.8° eccentricity had first fixations with higher standard deviations ($M=6.73\%$) compared targets at 10.6° eccentricity ($M=5.24\%$). Therefore, the scatter of the saccade landing position is greater when the target is closer to fixation. This result is the same as we found in Experiment 8. There was no main effect of Left-Right [$F(1,9)=0.014$, where $p=0.908$] or an interaction effect [$F(1,9)= 0.187$, where $p=0.676$].

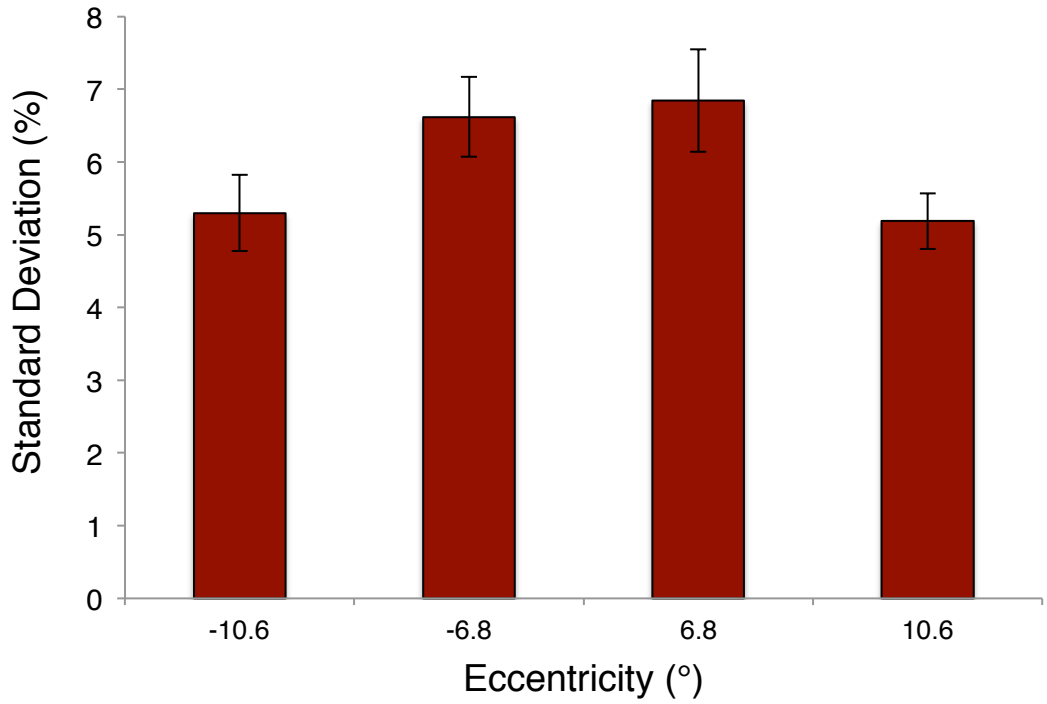


Figure 9.13. Average standard deviation of the first fixation as a function of target eccentricity. Negative (-) eccentricities represent targets to the left of fixation. Error bars represent the standard error of the mean.

9.2.2.3 Saccade Latency

The saccade latency (time to initiate a saccade) was analysed with a 2-way ANOVA (see Figure 9.14). A main effect of target position [$F(1,9)=55.909$, where $p<0.001$] was found, whereby Bonferroni corrected pairwise comparison shows that targets at 6.8° had shorter latencies (M=262ms) compared to targets at 10.6° eccentricity (M=303ms). This is in line with the findings of Experiment 8, where high eccentricity targets had longer latencies. There was no main effect of Left-Right [$F(1,9)=0.025$, where $p=0.877$] or an interaction effect [$F(1,9)= 1.229$, where $p=0.296$].

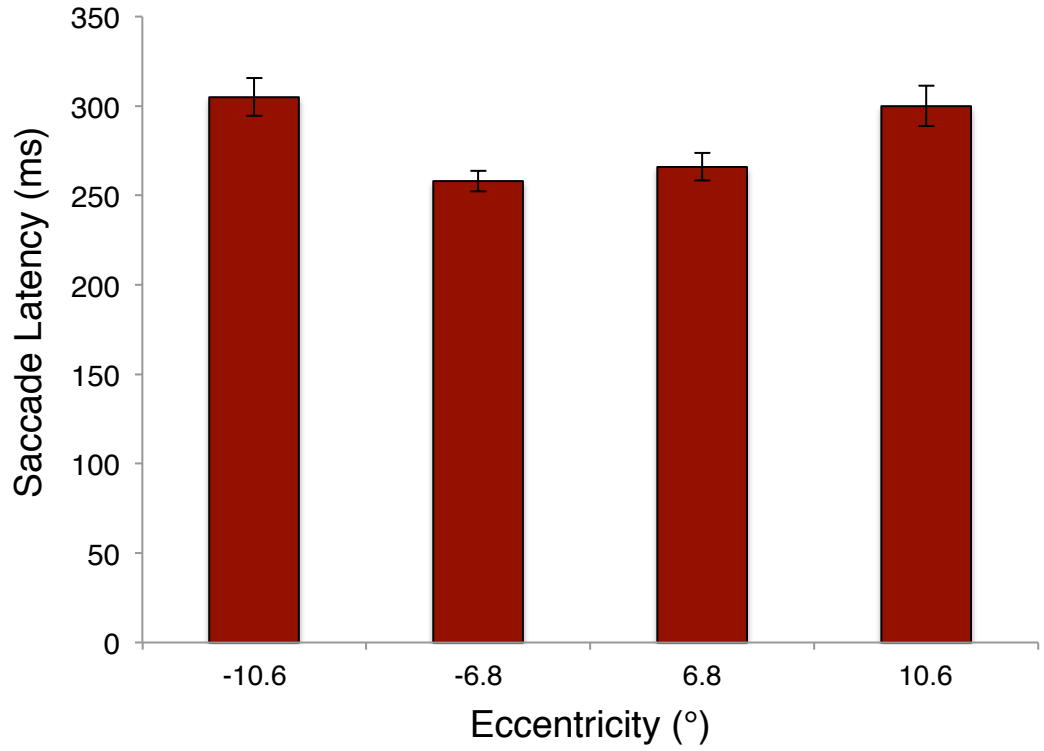


Figure 9.14. Average saccade latencies as a function of target eccentricity. Negative (-) eccentricities represent targets to the left of fixation. Error bars represent the standard error of the mean.

9.3 Experiment 10: Eye Movements to Triangle-Shaped Targets

From the results of the above experiments, we posit that saccades are being made to a reference point (the centre of the target) with a certain degree of undershoot. However, we do not know what this reference point is. Up until now, we have assumed that it is the centre of the figure. Here, we refer to the centre as the midpoint of the target width. However, there is also the geometric centre of the figure, which is the average position of all the points in the figure. We will refer to this point as the centre of gravity (COG), but it is also known as the centroid.

The centre and the COG of a figure may coincide with each other, especially with a symmetrically shaped object. This was the case for the rectangular shaped stimuli from Experiment 8, where both points (centre and COG) were identical to each other. Therefore, in this experiment, the shape of the target was a triangle – specifically, an isosceles triangle which has 2 sides with equal lengths. As a result, the COG and the centre of the figure do not coincide with each other (see Figure 9.16). This will allow us to determine whether a saccade is planned in relation to the target centre, or the target COG. In the case of the former, it implies that the landing position of the saccade is calculated based on the knowledge of the two lateral edge points of the figure border, while for the latter, it implies that the saccade position is calculated based on the knowledge of the entire figure shape.

9.3.1 Methods

9.3.1.1 Participants

10 participants (4 males) between the ages of 21 and 28 (M_{age} = 22.9) were recruited from the University of Nottingham Malaysia. All participants had normal or corrected-to-normal vision.

9.3.1.2 Apparatus & Display

The apparatus setup and display were the same as Experiment 8.

The stimuli presented were 1024×768 pixels, which was of a texture array that had a visual angle of 25.9°×19.5°. The triangular figure patch spanned 4.6°×7.6° (height × width), while the fixation point was 0.4° wide. The position of the centre of the figure patch varied at eccentricities of 6.8° (same as middle position of Experiment 8) and 8.4° to the left or right of a central fixation point.

9.3.1.3 Stimuli

The stimuli were pre-generated using PsychoPy 1.83.04 (Pierce, 2007) on an Apple Mac Mini. The overall texture array comprised 50 × 68 line elements, with a figure patch that was 12×20 elements (base × height). The figure patch could appear at four possible locations – at eccentricities of 6.8° and 8.4° to the left or right of fixation. This eccentricity is the distance between the central fixation cross and the centre of the target (mid point). Unlike the previous experiments that used rectangular shaped figures, the present experiment used triangular shaped figures (Figure 9.15) that that were pointing inwards or outwards relative to the central fixation.

For targets at the same eccentricity and side (i.e. left or right), the inwards and outwards pointing triangle will have the same centre point (half the height of the triangle). However, the COG of the outwards pointing triangle will be closer to fixation, while it is further away from fixation for the inwards pointing triangle (see Figure 9.16 for a schematic representation). At a viewing distance of 60cm, the distance between the centre and COG is 1.27°, while the distance between the COGs of the inwards and outwards triangle is 2.53°. The orientation of the line elements were constructed to form the Block profile (see Section

3.3), with orientation contrast of 90° between figure and ground elements. The orientation of the background elements (baseline orientation) was either 0° (vertical), 45°, 90°, and 135°.

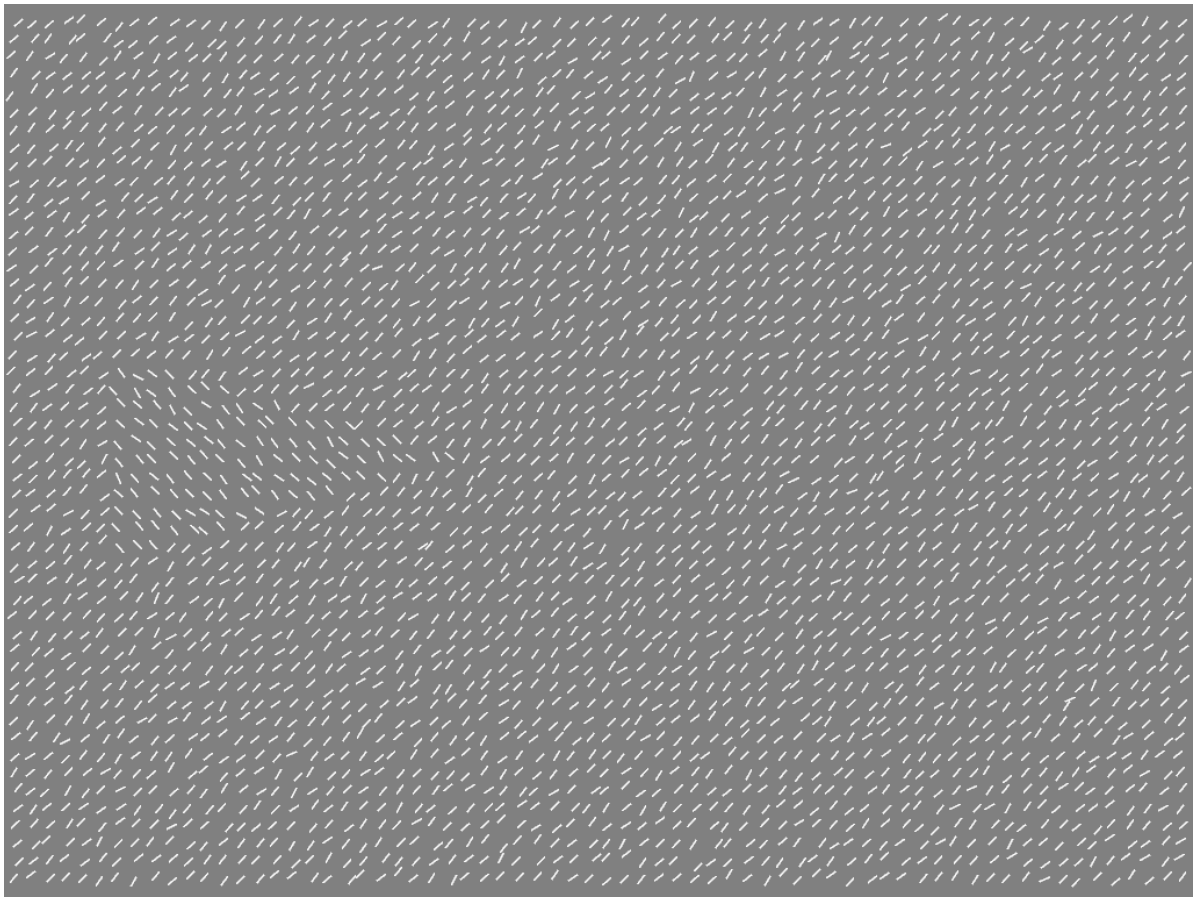


Figure 9.15. An inwards pointing triangular figure at 6.8° eccentricity to the left of fixation. The target figures are always of a Block profile with 90° orientation contrast. The targets in the experiment appeared inwards or outwards at eccentricities of 6.8 and 8.4° to the left or right of fixation.

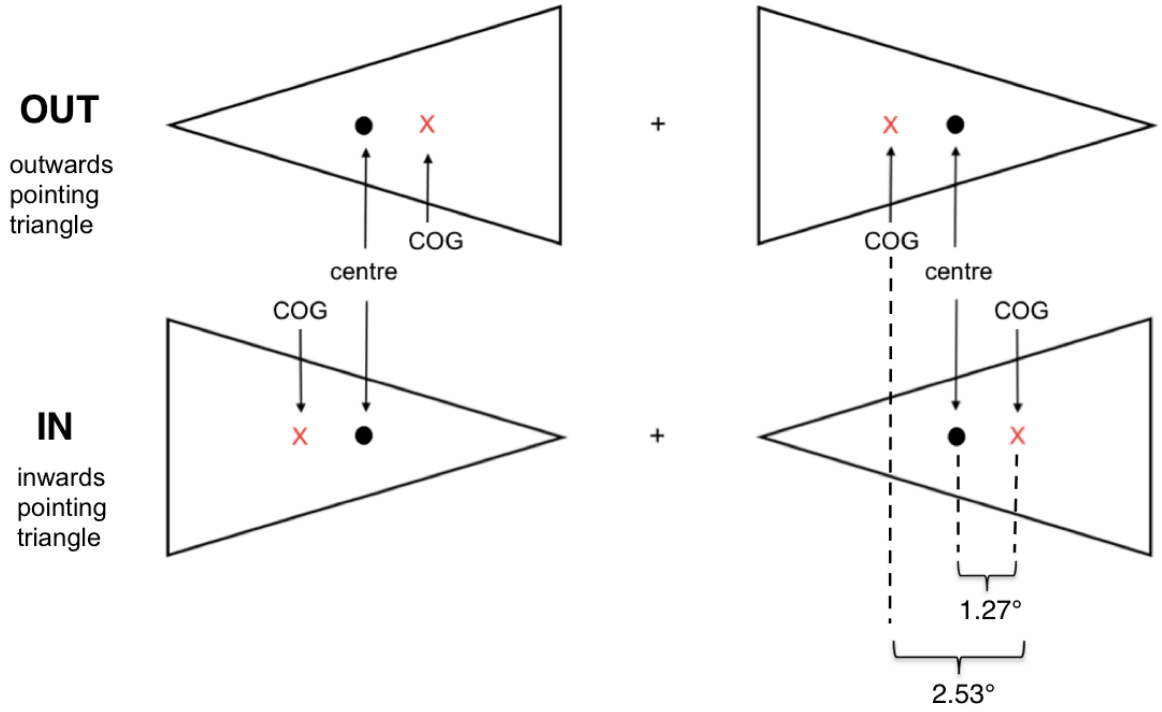


Figure 9.16. A schematic representation of the inwards and outwards pointing triangle. As can be seen from above, the centres (circular dot) of the two triangles (out vs. in) are positioned at the same spot, assuming the eccentricities are the same. The COG (×), is closer to fixation for the outwards triangle, while it further away for the inwards triangle.

9.3.1.4 Procedure

The instructions to the participants and the sequence of an experimental trial were the same as the Experiment 8 (Section 9.1.1.4). However, in the present study, participants performed 4 experimental blocks each (40 trials per block).

9.3.2 Results

Before the results were analysed, trials in which participants failed to make a primary fixation to the figure centre or edge (two columns on either side of the border) were removed.

9.3.2.1 Position of First Fixation

Once again the percentage amplitude (%) was calculated (see Section 9.1.2.1). In the current study, the eccentricities of the target centres were 6.8°and 8.4°. Saccade amplitude of 100% refers to the point at which a primary fixation was made to the figure centre (circular dot on Figure 9.16). For the target at 6.8° eccentricity, the amplitude distance to the edge closest to fixation is 45%, and the edge further from fixation is 155%. These amplitudes will be the same for both the inwards and outwards pointing triangle on both sides of fixation. However, the amplitude distance to the centre of gravity (COG) of the outwards pointing triangle is 82%, while it is a 119% for the inwards pointing triangle. For the target at 8.4° eccentricity, the amplitudes to the edges are 55% and 155%, while the COG of the outwards and inwards pointing triangle are respectively 85% and 115%.

Figures 9.17 and 9.18 show the mean amplitudes of the participants' first fixation, as well as the frequency histograms for the various conditions. The most apparent result we see is that the amplitudes for the inwards pointing triangle are not the same as the amplitudes for outwards pointing triangle. Because the saccade amplitude is calculated relative to the centre of the triangle (where amplitude to the centre is 100%), should participants be making a fixation to the centre or a fixed % of the centre, the amplitudes for both the inwards and outwards pointing triangle should be the same. We do not see this pattern of results, which suggest that participants' primary fixation is not to the centre of the figure.

In fact, the results show that amplitudes for the inwards pointing triangle is greater than the outwards pointing triangle. As a reminder, the COG for the outwards pointing triangle is 82% and 85% (6.8° and 8.4° eccentricity respectively), while the COG for the outwards pointing triangle is 115% and 119% (6.8° and 8.4° eccentricity respectively). Therefore, the fact that the inwards pointing triangle has amplitudes of primary fixations that are greater than the

outwards pointing triangle is an indication that the primary fixation is to the COG rather than the centre of the figure. Besides that, the histograms in Figure 9.18 show that the black and white bars (respectively inwards and outwards triangle) are spaced apart from each other, with the black bars (inwards triangle) having a greater number of high amplitude first fixations. This once again shows that the first fixation is being directed to the COG of the figure.

However, a key point to note is that the first fixation is not exactly to the COG, as there is still an undershoot. The black solid bars above each column in Figure 9.17 represent the amplitude of the COG, and inspection of the figure clearly shows that the amplitudes of participants' primary fixations are lower than the COG. Nevertheless, this is expected as we have consistently found in the previous experiments a certain degree of saccadic undershoot to the target. The undershoot to the inwards triangle (between 15 – 20%) is greater than the undershoot to the outwards triangle (~10%). A plausible explanation for this outcome is that compared to the inwards triangle, the COG for the outwards triangle is further away from fixation, and as we found in Experiment 9 – the further away the target is, the higher the undershoot is (poorer accuracy).

The amplitude data was also analysed using a 2 (left-right: target to left or right of fixation) × 3 (position: 8.4°, and 6.8°) × 2 (Out-In: triangles pointing outwards or inwards) ANOVA, the outcome of which is reported in Table 9.6.

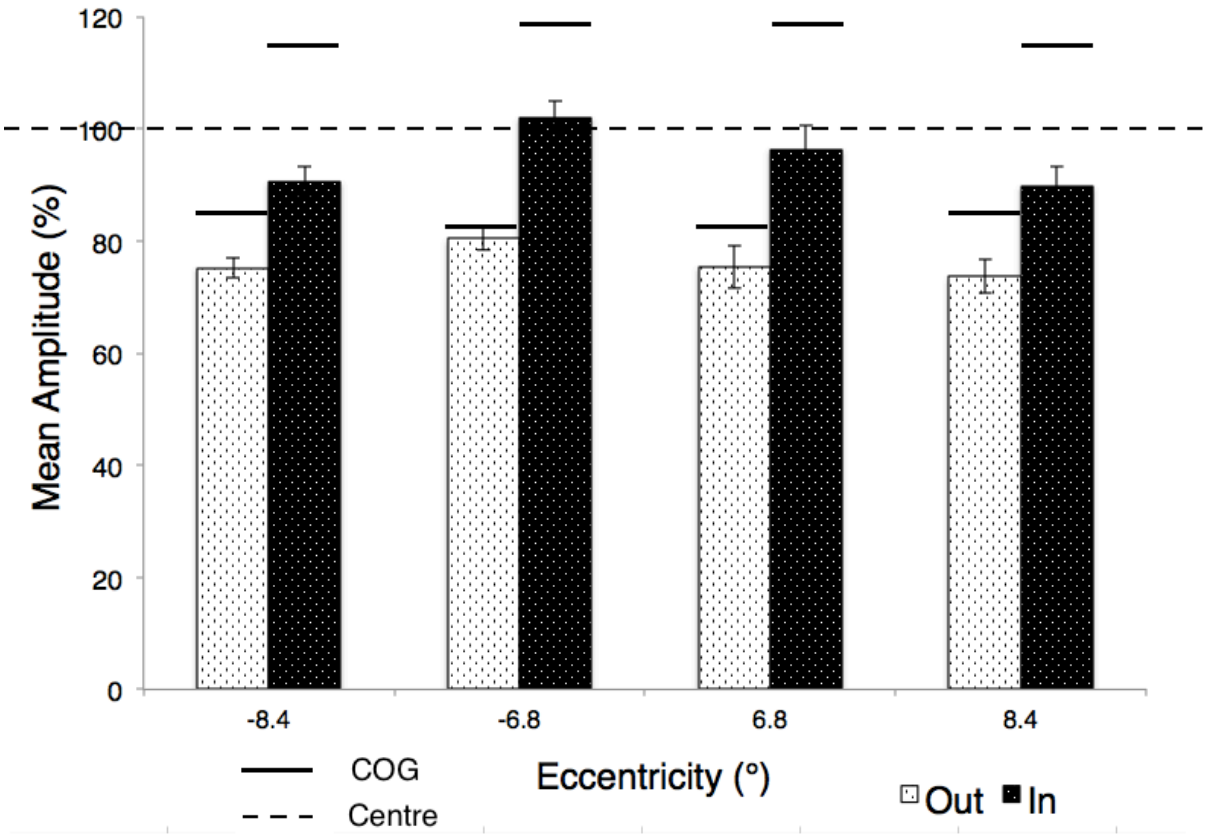
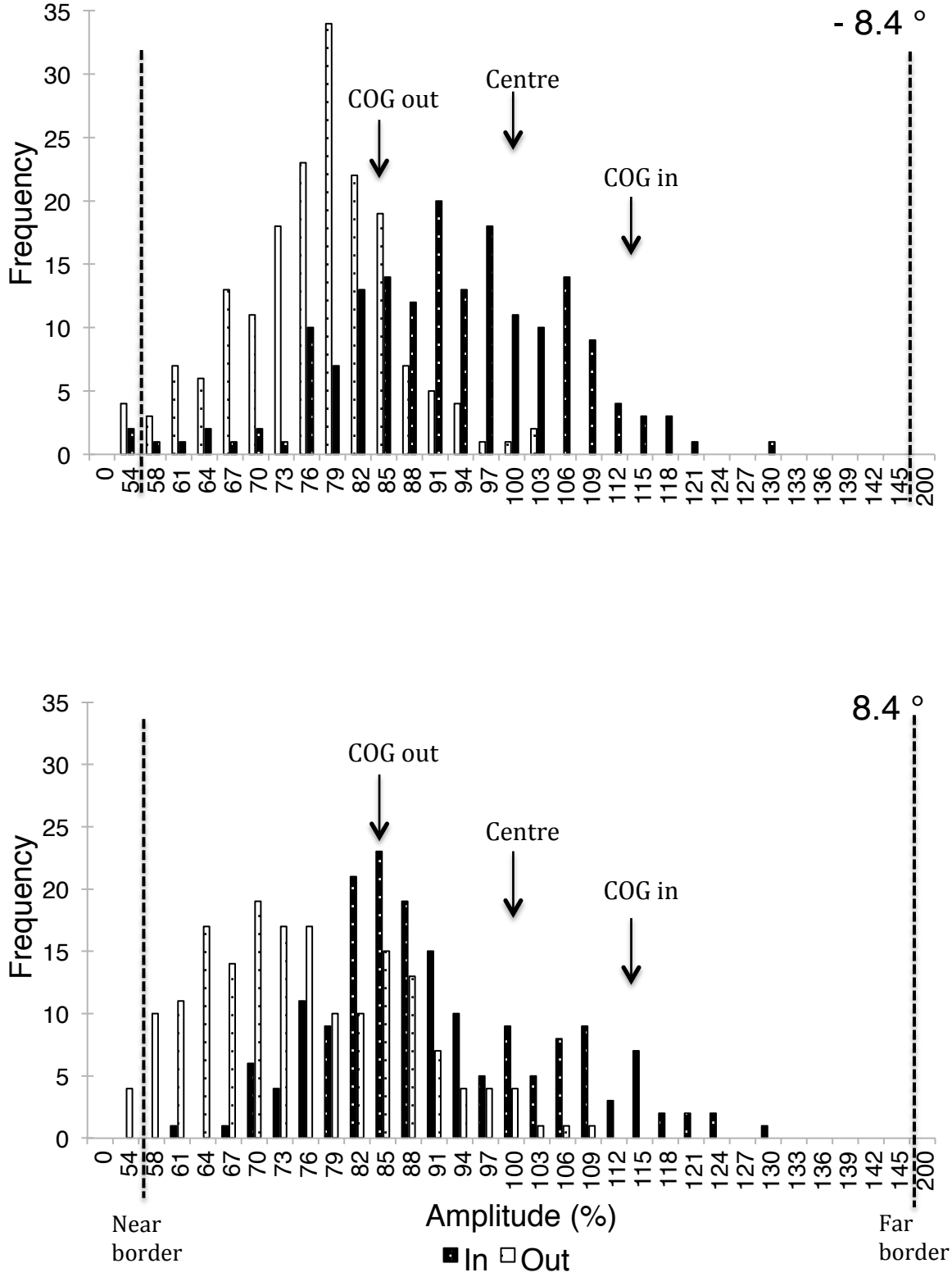


Figure 9.17. Average amplitudes of primary fixation in percentage relative to centre of figure (100%). The black solid bars above each column represent the COG for the individual conditions, while the dashed line is the centre of the figure. Negative (-) eccentricities represent targets to the left of fixation. *Out* is the amplitude for the outwards pointing triangle, and *In* is the amplitude for the inwards pointing triangle. Error bars represent the standard error of the mean.


- 8.4 °
COG out
Centre
COG in
Frequency
35
30
25
20
15
10
5
0
0 54 58 61 64 67 70 73 76 79 82 85 88 91 94 97 100 103 106 109 112 115 118 121 124 127 130 133 136 139 142 145 200
8.4 °
COG out
Centre
COG in
Frequency
Near border
Amplitude (%)
Far border
In
Out

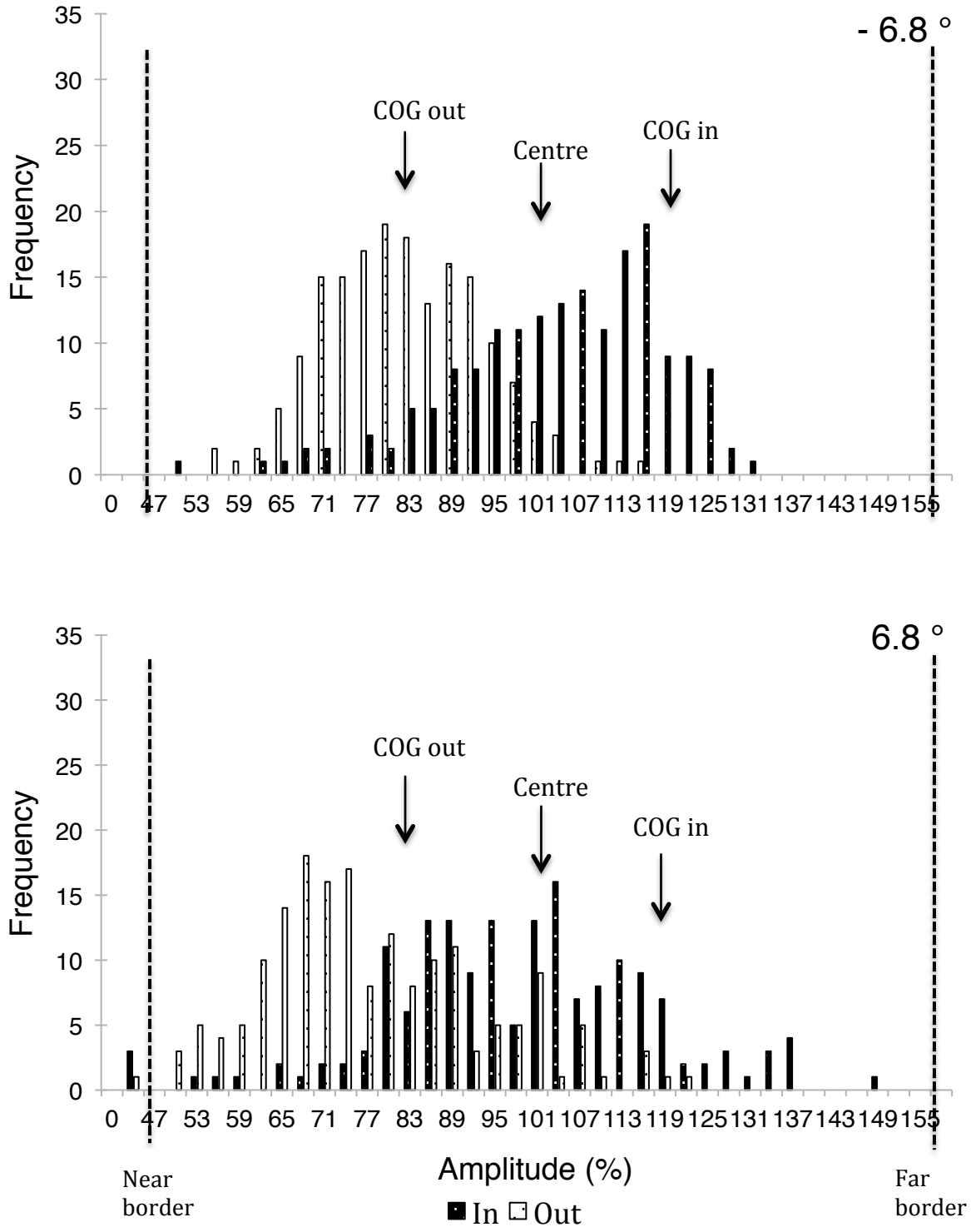


Figure 9.18. Saccade amplitude data (in percentage relative to figure centre) plotted as frequency histograms, with pooled data from all 10 participants. The direction and eccentricity of the target figure is indicated on the top right corner of the graph, and the dashed lines on the graph represent the border of the figure. *Out* is the amplitude for the outwards pointing triangle, and *In* is the amplitude for the inwards pointing triangle. The

intervals on the x-axis represent the upper limit of the bin array e.g. 83 includes amplitude up to 83%. The graph also indicates where the COGs (for both types of triangles) and the Centre is in amplitude. The spread seen between the two triangles is indication that the primary fixation is not to the centre (or we should see an identical pattern between the black and white bars).

Table 9.6
Summary of ANOVA results on Position of First Fixation (saccade amplitude in percentage relative to figure centre).

	df	*F*	*p*
Left-Right	1,9	1.349	0.275
Position	1,9	46.427	**<0.001**
Out-In	1,9	88.709	**<0.001**
Left-Right × Position	1,9	4.440	0.064
Left-Right × Out-In	1,9	0.000	0.991
Position × Out-In	1,9	14.215	**0.004**
Left-Right × Position × Out-In	1,9	0.520	0.489

A significant main effect of Position was found, whereby targets at 6.8° eccentricity (M=88.6%) had greater amplitudes than targets at 8.4° (M=82.3%) targets [where p<0.001 for Bonferroni corrected pairwise comparison]. This replicates the findings from Experiment 9, where saccades to targets with high eccentricity are less accurate (lower amplitudes). A significant main effect of Out-In was also found, which showed that outwards pointing triangles (M=76.2%) had lower amplitudes compared to inwards pointing triangles (M=94.7%) [where p<0.001 for Bonferroni corrected pairwise comparison]. The amplitudes to the centre of both types of triangles are identical at 100%. Therefore, this difference in amplitude of Out-In indicates that the primary fixation is being made to the COG, as the amplitude to the COG for the inwards triangle is larger than the outwards triangle.

Post hoc tests using Bonferroni correction was used to analyse the Position and Out-In interaction. The comparisons show that for the inwards pointing triangle, the position influences the amplitude greatly ($M_{6.8°}$=99.23%; $M_{8.4°}$=90.20%, p<0.001). However for the

outwards pointing triangle, the position only has a moderate effect on the amplitude ($M_{6.8°}$=77.90%; $M_{8.4°}$=74.40%, p<0.014).

9.3.2.2 Precision of Landing Position

The precision of the landing position was analysed with a 3-way (left-right × position × out-in) repeated measures ANOVA (see Figure 9.19 and Table 9.7). A significant main effect of Position was found, whereby Bonferroni corrected pairwise comparison shows that targets at 6.8° eccentricity had first fixations with greater variability (M=11.20%) compared targets at 8.4° eccentricity (M=8.43%). Thus, the reliability of the saccadic landing position is better when the target is further away from fixation (also found in Experiment 8 & 9). Furthermore, a significant main effect of Out-In was also found, in which pairwise comparison using Bonferroni correction showed that inwards pointing triangles (M=10.77%) had fixations with greater variability compared to outwards pointing triangles (M=8.87%). That is, saccades to the inwards pointing triangles were more scattered compared to the outwards pointing triangles. No other comparisons were significant.

Table 9.7
Summary of ANOVA results on Precision (Standard Deviation) data.

	df	*F*	*p*
Left-Right	1,9	0.002	0.967
Position	1,9	13.117	**0.006**
Out-In	1,9	11.301	**0.008**
Left-Right × Position	1,9	2.273	0.166
Left-Right × Out-In	1,9	0.125	0.732
Position × Out-In	1,9	0.428	0.530
Left-Right × Position × Out-In	1,9	2.604	0.141

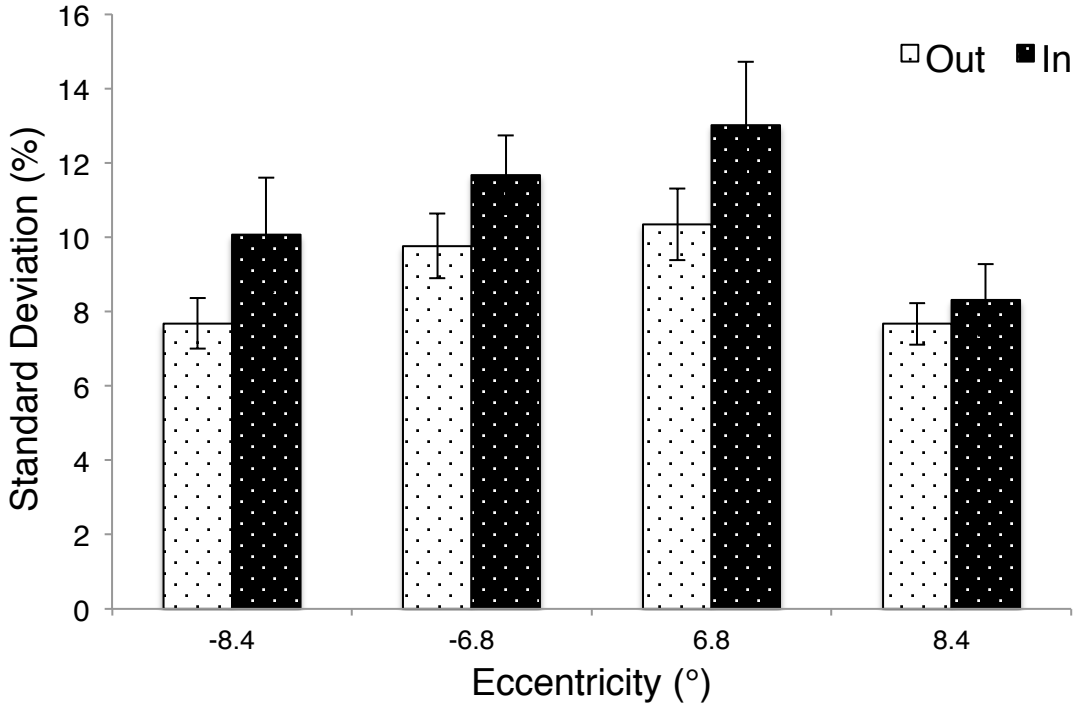


Figure 9.19. Average standard deviation for the outwards (*Out*) and inwards (*In*) pointing triangle as a function of target direction and eccentricity. Negative (-) eccentricities represent targets to the left of fixation. Error bars represent the standard error of the mean.

9.3.2.3 Saccade Latency

The saccade latency (time to initiate a saccade) was analysed with a 3-way ANOVA (see Figure 9.20 and Table 9.8). An inspection of the graph, as well as the outcome of the ANOVA shows that saccade latency did not vary with any of the conditions. In contrast with the other experiments (8 and 9), this study shows no increased latencies to high eccentricity targets (as there is no main effect of position). It is possible that this outcome is driven by the smaller difference in target eccentricities.

Table 9.8
Summary of ANOVA results on Saccade Latency data.

	df	*F*	*p*
Left-Right	1,9	1.855	0.206
Position	1,9	3.792	0.083
Out-In	1,9	0.239	0.636
Left-Right × Position	1,9	0.835	0.385
Left-Right × Out-In	1,9	1.173	0.307
Position × Out-In	1,9	0.009	0.927
Left-Right × Position × Out-In	1,9	4.327	0.067

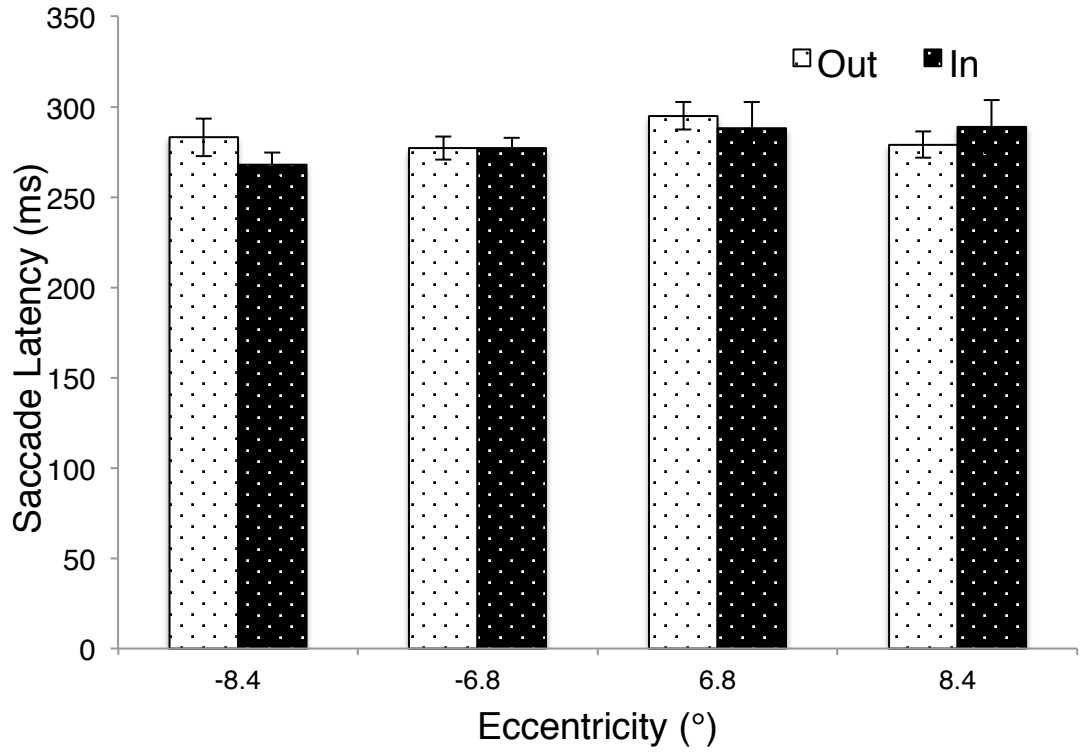


Figure 9.20. Average saccade latency for the outwards (*Out*) and inwards (*In*) pointing triangle as a function of target direction and eccentricity. Negative (-) eccentricities represent targets to the left of fixation. Error bars represent the standard error of the mean.

9.4 Discussion

A key finding of Experiment 8 was that participants were able to localise the figure patch. At all figure eccentricities, primary saccades were made to the figure location. Unlike Experiment 7 (presented in Chapter 8), the results of the current study are in line with studies (e.g. Deubel et al., 1988; Deubel & Frank, 1991; Supèr, Spekreijse, & Lamme, 2001) that show it is possible to localise the figure location and make a primary saccade to the figure. Interestingly, when the saccade landing position was calculated as percentage amplitude, the amplitude was similar for all eccentricities – approximately 85%. This undershoot to the target position is consistent with findings in the literature that show saccades tend to undershoot their targets (e.g. Deubel et al., 1988; Findlay, 1982; Land & Tatler, 2009). However, the 15% undershoot found in this study was higher than the 5-10% undershoot found by Deubel et al., 1988 with their texture stimuli.

Unlike our findings, Kapoula and Robinson (1986) found that participants were able to make accurate saccades i.e. without the undershoot, only for near targets (red spot, and E/∃) at 5° eccentricity. Overall, our results do not show this range-effect, whereby saccades overshoot near targets and undershoot far targets. In Experiment 8, we had targets at 4.5° eccentricity, but there was still an undershoot of 15% regardless of the proximity to fixation. This difference in results may well have been due to the different sizes of the targets used. The size of the targets in Experiment 8 were 5.7°×4.6°, while it was 0.4°×0.2° in Kapoula and Robinson's experiment, and studies have shown that smaller targets are more accurate (Kowler & Blaser, 1995; Ploner et al., 2004).

In fact, target width was 2° (Deubel & Frank, 1991) and 2.1° (Deubel et al., 1988) compared to our 4.6°, which explains why our experiments show larger saccadic undershoot compared

to theirs. Moreover, the notion that smaller targets are more accurate is corroborated by our findings from Experiment 9, where we had targets that were smaller. In Experiment 8, the width of the target was 4.6°, while in Experiment 9, the width of the target was 1.5°. With the larger targets, the amplitude percentage was between 80 – 90%, while for the small targets the amplitude was between 94 – 98%. Additionally, the scatter of saccades was greater with larger targets (7 – 17%) compared to smaller target (5 – 7%), thus showing not only are smaller saccade more accurate, but they are also more precise.

The initial aim of Experiment 9, where we decreased the size and increased the eccentricity of the target, was to investigate if saccades will have greater amplitudes to ensure that the landing position of the first fixation fell within the figure border. Despite the fact that this was the outcome we obtained, we cannot conclusively claim that the increased saccade amplitudes to these targets was a result of the visual system accounting for the increased amplitude needed to program a saccade that fell within the figure border. This is because saccade amplitudes of 85% would fall within the figure border of the 6.8° eccentricity target in Experiment 9. Yet, we found that saccade amplitudes were on average 98.7%. This is 13.9% higher than the saccade amplitude for the 6.8° eccentricity target in Experiment 8, where the figure size was 3× bigger. Thus, the higher saccade amplitude in Experiment was driven by the smaller size of the figure (as discussed in the previous paragraph).

However, inspection of Figure 9.12 shows that some saccades did land outside the figure border. For the targets at 6.8° eccentricity, approximately 8% of trials fell outside the figure-defined border. On the other hand, 29% of trials fell outside the figure border for the target at 10.6° eccentricity. As a reminder, the amplitude distance to the border closest to fixation is 93% for the 10.6° target and 89% for the 6.8°target. So while the overall mean amplitude

data shows that participants made saccades to the target figure, a greater number of trials still fell short of actually reaching the confines of the figure border.

Experiment 8 also reveals that the saccadic landing position is not influenced by the different edge profiles. Only at the closest eccentricity (4.5°) was there an effect of profile, where saccades to the Cornsweet profile had lower amplitudes compared to the Block and Blur profiles. According to the salience-based model of eye movements developed by Itti and Koch (2000), the destination of a saccade is programmed based on a highly salient part of a scene. In their model, early visual features (orientation, luminance, colour) are processed pre-attentively, and a winner-take-all mechanism detects a location that is most salient and directs attention to that location. The Cornsweet profile has information of orientation contrast at the edge of the figure only, making that the most salient part of the stimulus. Therefore, one could suggest that the lower amplitude of the saccade to the Cornsweet profile is due to the saliency of the figure at the edge region. However, the amplitude distance to the figure border (edge) was 50%, and the saccade amplitude to the Cornsweet profile was 82%, which is closer to figure centre than figure edge. Thus, the salience-driven model of eye movements cannot explain our results.

In fact, our findings are more in line with that of Melcher and Kowler (1999) who suggest that saccades are determined after the integration of information to segment the figure from the ground. As such, the overall shape of the figure guides the eye movements, not just the salient parts. This explains why the different profiles do not have an impact on the landing position of the primary fixation as all profiles have the same shape. Additionally, the results of Experiment 10 show that saccades are programmed **not** to the centre of the figure (i.e. middle position), but to the centre of gravity (COG). If the saccade destination was the centre of the figure, the saccade amplitude to the outwards and inwards pointing triangle should

have been the same. However, our results show that there was a difference in saccade amplitude of at least 15% between the two triangles.

We also found that the undershoot to the inwards triangle (between 15 – 20%) is greater than the undershoot to the outwards triangle (~10%). This is possible due to the COG of the inwards triangle being further away than the outwards triangle for a particular eccentricity. Studies have shown that targets at greater eccentricities have larger undershoot than targets at nearer eccentricities (e.g. Ploner et al., 2004; Kapoula & Robinson, 1986). We have shown the same effect in Experiments 9 and 10, where saccadic undershoot was greater for further targets. These results can be interpreted as the spatial pooling of information being more complex for targets further away, which results in a decreased accuracy of the saccadic landing position.

Additionally, we found a difference in saccade latencies for targets that were further away (Experiment 8 and 9), which is expected as the further away the target from the fovea, the longer the saccade latencies (Land & Tatler, 2009; Gilchrist, 2011). We also found that the time to initiate a saccade to the Blur profile was longer than the Block or Cornsweet profiles. This outcome is probably due to targets with greater intensity/saliency eliciting shorter latencies (Sumner, 2011; Gilchrist, 2011; Marino & Munoz, 2009). According to Supèr, Spekreijse, and Lamme (2003), the stronger the contextual modulation (stronger neuronal response to figure then ground), the shorter the saccade latency. They propose that the visual system uses neural activity in V1 as internal evidence of a stimulus to guide saccades. This occurs when a neural response is made by the sensory system after stimulus onset. A build-up in the motor system is produced prior to saccade onset. When the strength of the figure-ground signal is great, the build-up in the motor system will be rapid, which in turn produces faster saccades (Supèr & Lamme, 2007). This implies that when a stimulus is not very

salient, the neuronal activity will be weak, and saccade latencies will be longer. The Blur profile is the least salient amongst the three profiles – evident by psychophysics (Experiments 1 – 6), and eye tracking (Experiment 11 in Chapter 10) results. Therefore, increased saccade latencies for the Blur profile could be accounted by the low saliency of the Blur profile that produced weaker signal strength, which in turn produced longer latencies. This finding also lends support to the idea that a figure boundary is formed (i.e. a region is segmented) **before** a saccade is planned as suggested by Melcher and Kowler (1999), and Deubel and Frank (1991).

Finally, the key outcome from the area-normalized fixation count and summed fixation duration measure is that participants on average made more and longer fixations to the centre region than the edge region. This includes the Cornsweet profile, where orientation contrast information is present only at the edge. This is once again in contradiction to the salience accounts of eye movements (e.g. Itti & Koch, 2000; Parkhurst et al., 2002) as participants looked longer and more frequently at a region of low salience. However, this pattern of results could potentially be explained by viewing the figure as a whole, along with the tendency to fixate on the centre of the object (Tatler, 2007).

Chapter 10

Eye Movements Driven by High Saliency Stimuli with Orientation Contrast at the Edge and Centre of Figure

10.1 Experiment 11: Matched Saliency Task

In the experiments presented in Chapters 8 and 9, participants were shown texture stimuli, and had to **detect** the figure embedded within the texture grid. Overall, the findings show that orientation contrast and orientation profile had a marginal effect on the total fixation duration and numbers of fixations made to the centre and edge region. Furthermore, the type of orientation profile did not affect the landing position of the first fixation, but it did affect the saccade latency, whereby the time taken to initiate a saccade to the Blur profile is increased compared to the Block and Cornsweet profiles. This difference in time taken to initiate an eye movement to the different profiles suggests that the mechanisms that segregate the textures are being affected by the different profiles. However, all the previous studies conducted either had fixed orientation contrast (at 90°) or orientation contrast that was matched to a naïve observer's threshold from a psychophysics experiment. While this allowed us to draw conclusions about eye movements to textures with either high or low orientation contrast, it does not directly inform us about the saliency (quality of being particularly noticeable to stand out from others in its surrounding) of the different orientation profiles. We thus aim to investigate the subjective saliency of various texture profiles to test the role texture edges play in guiding eye movements.

In the current experiment, two texture patterns were presented simultaneously to participants on either side of a central fixation point. Participants had to make a saccade to the texture that

had an embedded figure that was more detectable. Besides manipulating the orientation contrast of the texture, we manipulated the type of orientation profile presented to the participants (i.e. different profiles will be presented on either side of fixation). This study will allow us to determine what values of orientation contrast will produce **matched** saliencies between the different profiles, which in turn will tell us which profile produces the highest levels of saliency to drive eye movements. That is, had we used a non eye movement measure (for example, by running this experiment as a psychophysical experiment), we would not be able to claim that specific profiles are best suited to drive eye movements.

If orientation contrasts at **both** the centre and edge of a figure are required to produce the highest levels of saliency for driving eye movements, it is expected that the values of orientation contrast needed by the Blur and Cornsweet profiles to match the Block profile will be significantly higher. If orientation contrast at only the centre of the figure was sufficient to drive eye movements, we would expect that the degree of orientation contrast needed to match Block profile would be equivalent for the Blur profile, and significantly higher for the Cornsweet profile. Likewise, if eye movements are driven by the properties of orientation contrast at the edge of the figure, orientation contrast values would be similar for both the Block and Cornsweet profiles, but the Blur profile would require higher levels of orientation contrast to have equal saliency.

10.1.1 Methods

10.1.1.1 Participants

10 participants (M_{age} = 21.8, SD_{age} = 1.1, 1 Male) were recruited from the University of Nottingham Malaysia. All participants had normal, or corrected-to-normal vision

10.1.1.2 Apparatus & Display

A Tobii T60 Eye-Tracker running Tobii Studio 3.2 was used to present the stimuli and record participants' eye movements. Participants sat 60cm away from the monitor, which produced a resolution of 1.52 arcminute per pixel. Each participant performed a nine-point calibration before the experimental task. Analyses of eye movements were done using the built-in functions of Tobii Studio, which worked in conjunction with Microsoft Excel to export the data for further statistical analysis.

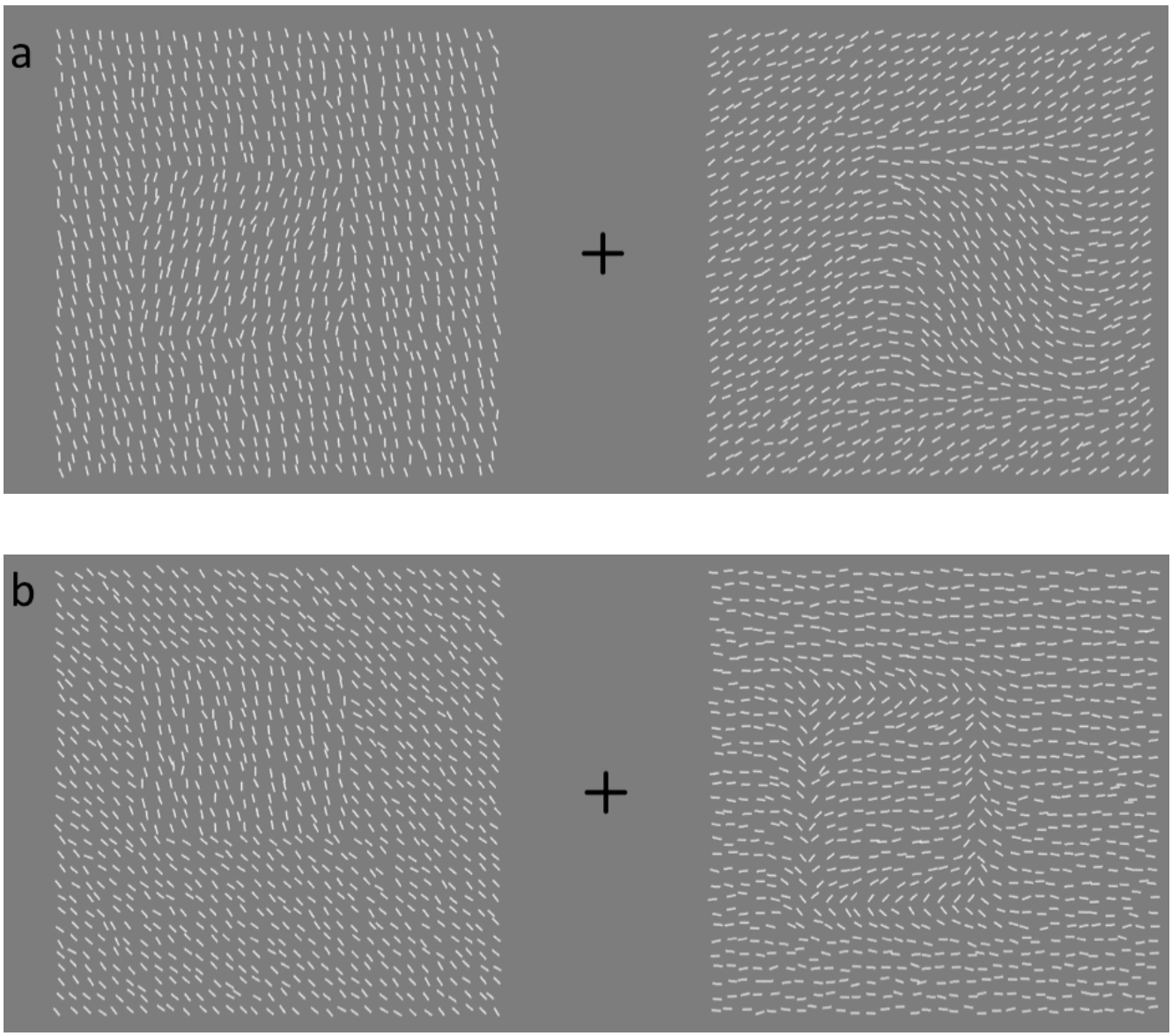
The screen size was 1280×900 pixels, which had a visual angle of 32.4°×22.8°. Each of the texture grids (on either side of fixation) had a visual angle of 12.2°, with the figure patch having visual angles of 5.7°×4.6° or vice versa. The fixation cross in the centre spanned 1.1° in visual angle, and the eccentricity to the centre of the texture grid was 8.9°.

10.1.1.3 Stimuli

The stimuli were pre-generated using PsychoPy 1.83.04 (Pierce, 2007) on an Apple Mac Mini. They comprised two texture grids presented on either side of a central fixation cross (see Figure 10.1). The texture grids were the same as described in Section 3.3.

Four conditions were tested: Block vs. Blur, Block vs. Cornsweet, Cornsweet vs. Blur, and Uniform for catch trials (Figure 10.1). To generate psychometric functions for each condition, one profile had a constant orientation contrast, while the other had orientation contrast that varied between 10° and 90°. For the varying texture, the figure was undetectable when orientation contrast was low, and extremely salient when orientation contrast was high (see Figure 10.2). The psychometric plots use the orientation contrast of the varying profile on the x-axis, and the proportion of "varying profile chosen" responses on the y-axis, and thus create a sigmoidal shape ranging from 0 to 100%. There were 5 different orientation contrast values used for the varying profile: 10°, 30°, 50°, 70°, and 90°. For the constant

profile, the orientation contrast was set to 30° for the Block (Block vs. Blur & Block vs. Cornsweet condition) and 50° for the Cornsweet (Cornsweet vs. Blur condition). These were determined by piloting the study beforehand.



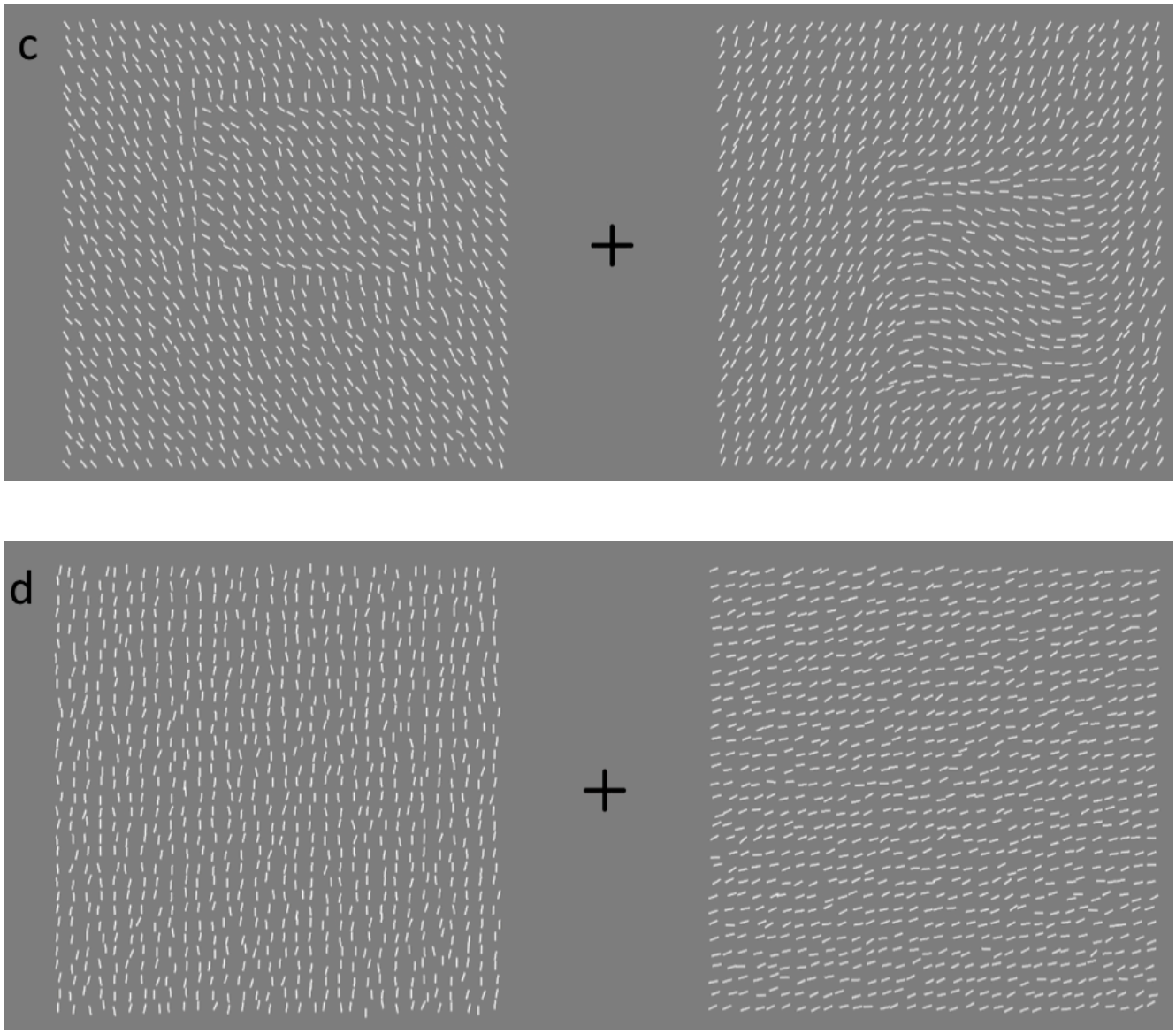


Figure 10.1. Example stimuli for the four conditions. (a) Block Vs. Blur: Block=30°, Blur=90° (b) Block Vs. Cornsweet: Block=30°, Cornsweet =90° (c) Cornsweet Vs. Blur: Cornsweet =50°, Blur=90°, and (d) Uniform condition with no figure present in either texture grid.

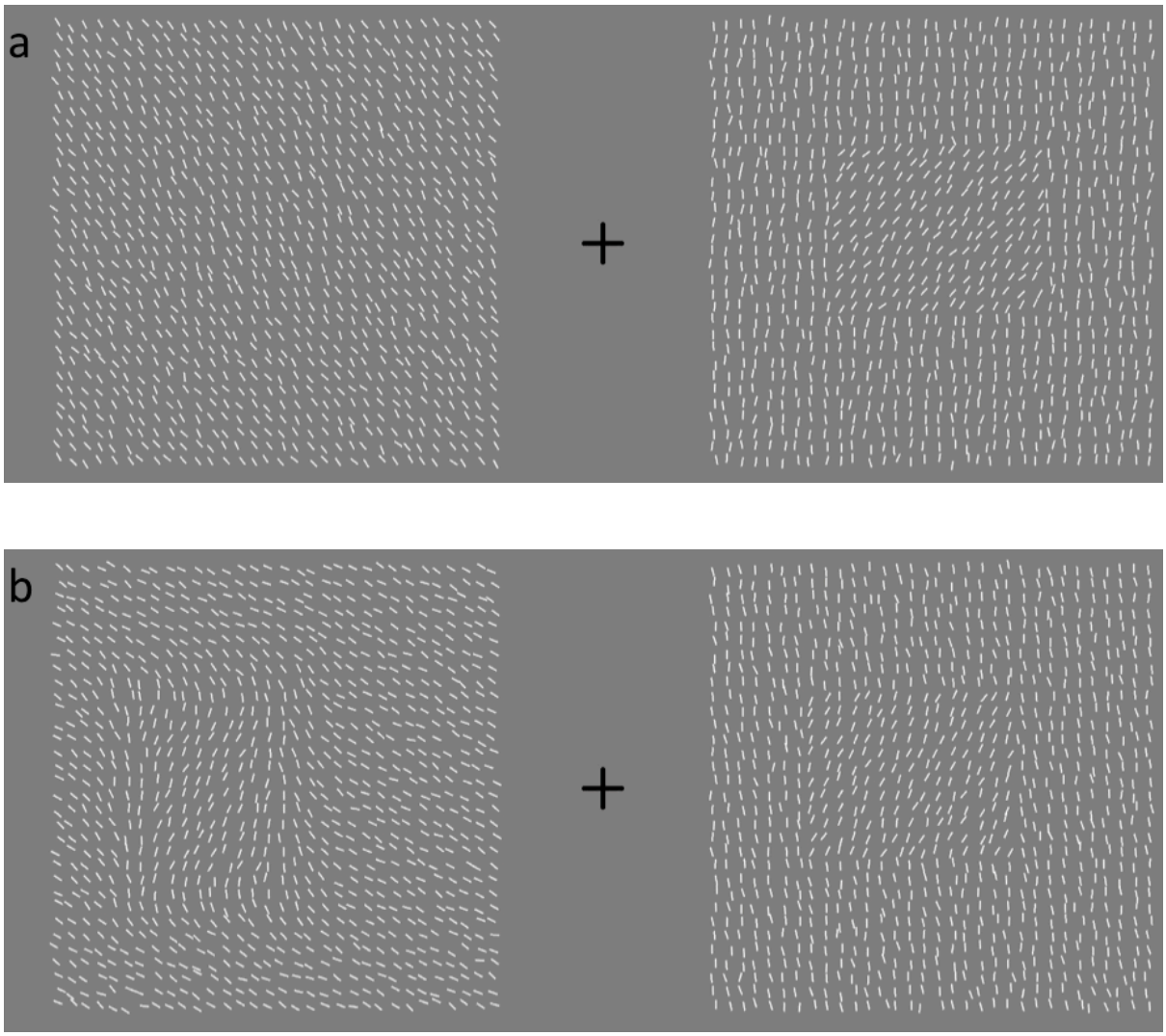


Figure 10.2. Example stimuli of the Block Vs. Blur condition. (a) Blur=10°, Block=30° and (b) Blur=90°, Block=30°. The Blur profile, to the right of fixation, is barely detectable in Figure 10.2a, but is much more detectable in 10.2b.

10.1.1.4 Procedure

There were 10 experimental runs in total, with each experimental run having 75 stimuli randomly presented from the four conditions. Before beginning the experimental run, participants performed several blocks of practice trials to familiarize themselves with the experiment.

A logistic distribution function was used (see Section 3.1.4) to determine the value of orientation contrast of the varying profile that has equal saliency to the constant profile. There were 750 trials in total, 150 (20%) of which were Uniform catch trials, and 200 each for the Block vs. Blur, Block vs. Cornsweet, & Cornsweet vs. Blur. All these trials were of different stimuli i.e. no stimuli were repeated. The method of constant stimuli was used. For each condition, the 200 trials are made up of equal number of trials for each value of orientation contrast (varying profile).

The fixation cross was displayed for 1 second, after which the stimulus patterns were displayed for 1 second. Following that, a new trial was initiated with the presentation of a fixation cross once again. Participants were instructed to fixate at the cross in the middle of the screen and move their eyes to the texture grid that had a figure that was most noticeable when the stimulus appeared. For catch trials, participants were required to maintain fixation after stimulus onset. Specifically, the instructions were: *"Your task is to look at the texture in which the embedded figure is more noticeable. Once you have moved your eyes to either one of the textures with a more noticeable figure, maintain you gaze until the image disappears. Once the image has disappeared, move your eyes to back to the fixation cross. If you find none of the textures contain a figure embedded in it, do not move you eyes, and maintain your gaze at the fixation cross."*

10.1.2 Results

On some trials (4.48%), participants made fast subsequent saccades from one texture grid to the other. That is, after stimulus onset, an eye movement will be made to one of the texture grids, immediately followed by another eye movement to the opposite texture grid. These trials were removed before analysis. On average, participants were able to make a saccade in 90.7% of the figure-present trials. For the remainder 9.3% of the figure present trials, they

maintained fixation at the fixation cross. These trials were not analysed as participants' response of maintaining fixation indicated that they did not perceive a figure within either texture grid.

For each participant, three psychometric functions were plotted (1 for each condition: Block vs. Blur, Block vs. Cornsweet, & Cornsweet vs. Blur). The y-axis of the graph was the percentage varying profile chosen while the x-axis was orientation contrast of the varying profile. KaleidaGraph was used to produce a curve of best fit (logistic distribution function, see section 3.1.4).

The curve fitting function was used to obtain the Point of Subjective Equality (PSE), which informs us of the orientation contrast needed by the varying profile to match the saliency of the constant profile. E.g. for the Block (constant) vs. Blur (varying) condition, the PSE value is the orientation contrast needed by the Blur profile to match the saliency of the Block at 30° orientation contrast.

A total of 30 psychometric functions were plotted (10 participants, with 3 conditions each, see Figure 10.3 for an example), with 30 PSE values generated. For each condition, a one-sampled t-test was used to analyse if the PSE values were significantly different from the test value, which was the orientation contrast of the constant profile.

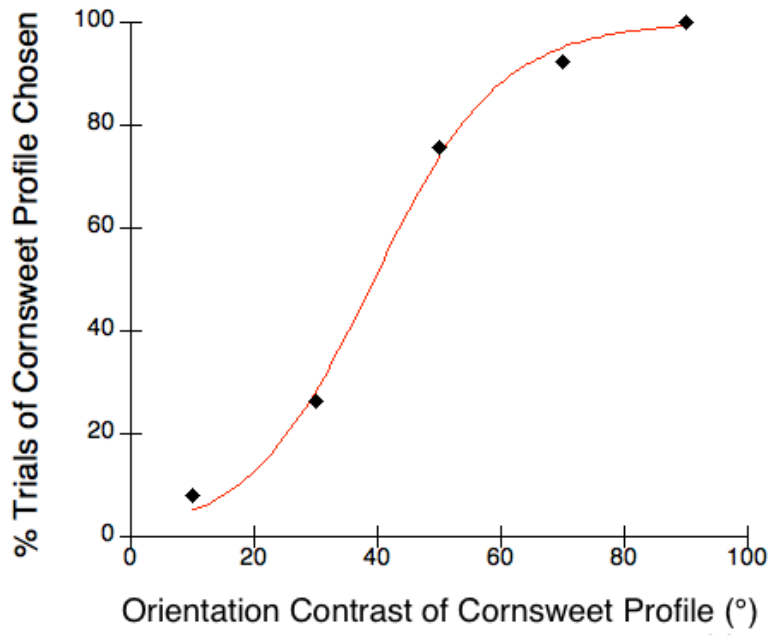


Figure 10.3. Example of one psychometric plot for one participant (Cornsweet vs. Block condition). The PSE for this plot is 39.49°, which is the orientation contrast needed by the Cornsweet profile to match the saliency of the Block profile (at 30° orientation contrast).

For the Block vs. Blur condition, mean PSE value (M=66.53°, $S.E.M$=2.12°) was significantly higher than the orientation contrast of the Block profile of 30°, $t(9)$=17.251, p<0.001. Likewise, for the Block vs. Cornsweet condition, the PSE value (M=39.74°, $S.E.M$=0. 57°) was greater than the 30° orientation contrast of the Block profile, $t(9)$=17.139, p<0.001. Finally, for the Cornsweet vs. Blur condition, the average PSE value (M=103.88°, $S.E.M$=3.23°) was significantly greater compared to the contrast of 50° for the Cornsweet profile $t(9)$=16.705, p<0.001.

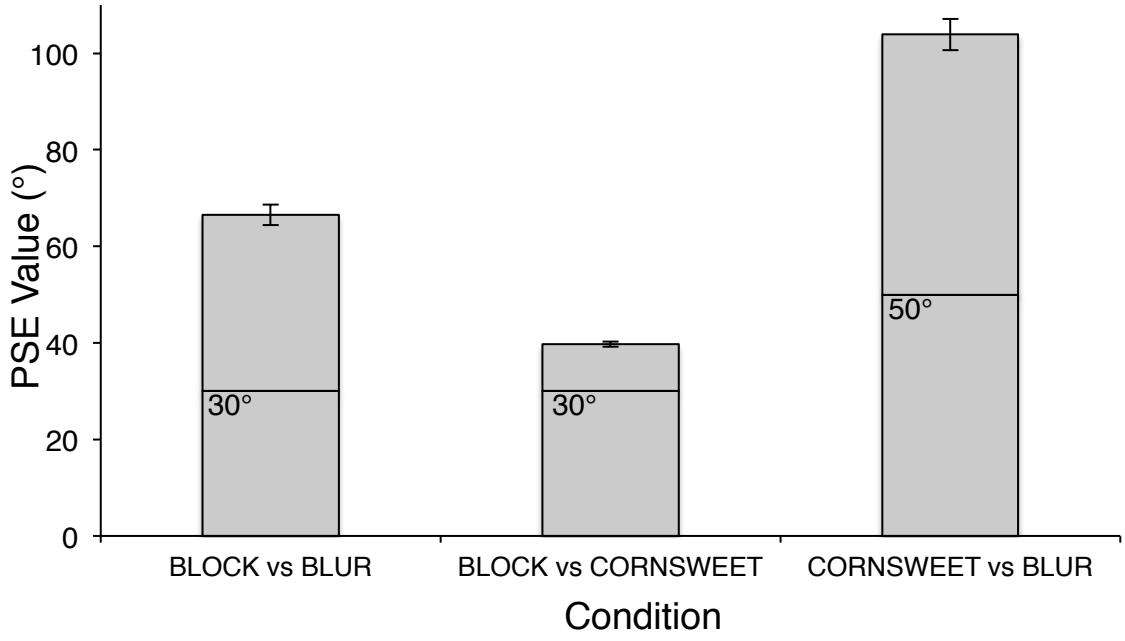


Figure 10.4. The graph shows the PSE values for each condition. The PSE values represent the orientation contrast of the varying profile to have equal saliency to the constant profile (black line on the bars). Error bars represent the standard error of the mean.

For the Cornsweet vs. Blur condition, the results indicate that when the orientation contrast of the Cornsweet profile is 50°, the orientation contrast of the Blur profile would have to be approximately 103° to have equal saliency. 103° orientation contrast is a value that is not physically possible as the maximum orientation contrast that can be achieved is 90°. However, the results nevertheless indicate that higher orientation contrast is needed by the Blur profile to match the saliency of the Cornsweet profile. A further paired-sample t-test confirmed this, as it was found that the average PSE value for the Block vs. Blur condition was significantly higher than the average PSE value for the Block vs. Cornsweet condition, $t(9)=11.329$, $p<0.001$. That is, to match the saliency of the Block profile at 30° orientation contrast, the orientation contrast of the Cornsweet profile had to be higher at 39°, and much higher at 66° for the Blur profile. Thus, the orientation profiles show a clear hierarchy of saliency, whereby Block > Cornsweet > Blur.

10.1.3 Modeling

We measured the levels of orientation contrast needed by the three profiles (Block, Blur, & Cornsweet) to have equal levels of saliency. We found that the Block profile was the most salient orientation profile, followed by the Cornsweet profile, then the Blur profile. As in the psychophysics experiments (Chapters 4 – 6), we have orientation thresholds for the three profiles, which allowed us to model the results (see Section 3.5). Briefly, the rationale of modeling the results is because the extraction of edge information by the visual system may well extend over a region beyond the immediate border of the figure i.e. not just one column on either side of border, but possibly more than one column. We proposed that a common mechanism averages the orientation on either side of the border for all three profiles. The size over which the mechanism operates is the Integration Region (IR), which is the point at which the orientation thresholds are most similar between the three profiles.

To model the results, we used each participant's PSE values for the Blur and Cornsweet profiles to calculate their individual size of integration region (IR, see Figure 10.5). The orientation contrast for the Block profile was always 30°. The IRs for the participants were between 2.34 and 3.17 (M= 2.77; SD=0.24). This is consistent with the size of the IR we found in Experiment 1, which was on average 2.5 for the segmentation task, and 3.5 for the detection task. The similar IR for the psychophysics experiment and this experiment implies that the saccadic system has access to the information produced by a mechanism that segregates figure-ground texture.

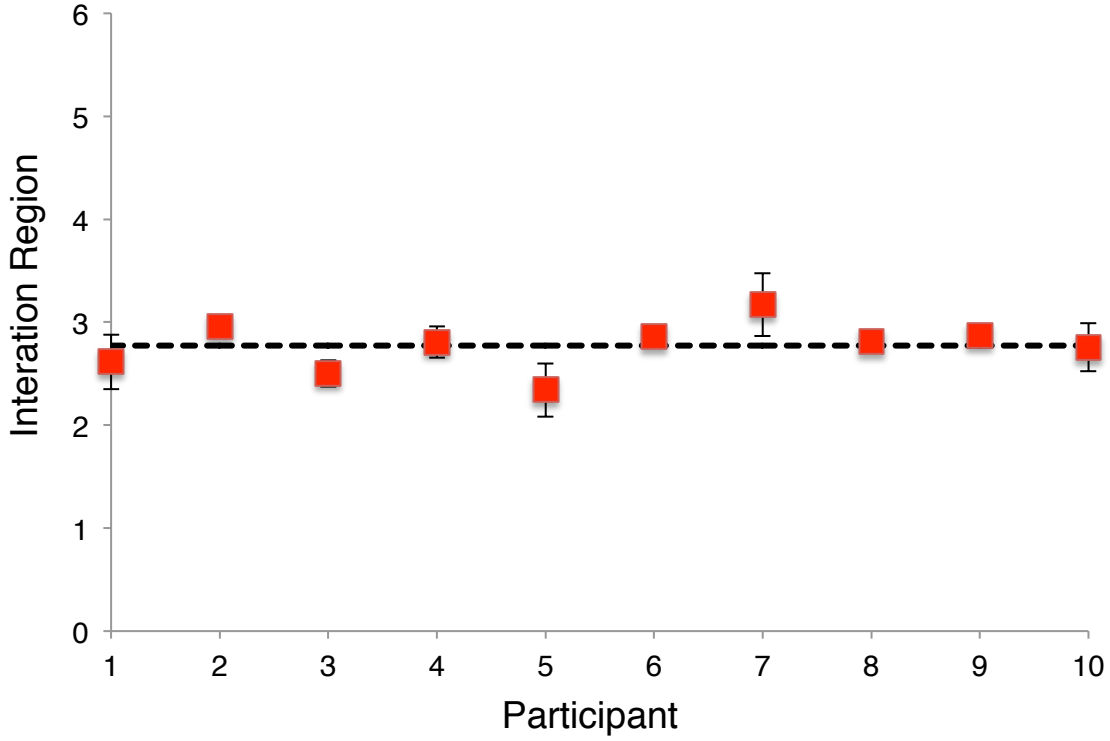


Figure 10.5. The graph above shows the integration region for all of the 10 participants, with the dashed line showing the mean integration region. The error bars represent the standard deviation, which was derived by bootstrapping the data.

10.2 Discussion

In the current study, PSE values were determined for each participant and condition. The mean PSE values were the mean orientation contrast needed for the varying profile to match the saliency of the constant profile. The three profiles were constructed to have different information of orientation contrast at the figure edge and centre. To elaborate, the Block and the Blur profile contained similar information of orientation contrast at the centre of the figure, and the Block and the Cornsweet profile contained similar information of orientation contrast at the edge of the figure. Therefore, if the PSE value for the Blur profile was similar to the orientation contrast of the Block profile (30°), it would suggest that the centre region was driving the eye movements. Likewise, if the PSE value for the Cornsweet profile was

similar to the orientation contrast of the Block profile, it would suggest that the edge region was driving the eye movements.

However, our findings here show that the PSE values for both the Blur and Cornsweet profile were higher than the orientation contrast of the Block profile. These results as a whole imply that orientation contrasts at both the edge and centre of the figure are required to produce the highest levels of saliency for driving eye movements. Furthermore, the Cornsweet profile was more salient than the Blur profile because of the following two results. One, the orientation contrast needed by the Blur profile to match the saliency of the Cornsweet profile (at 50° orientation contrast) was higher. Two, a comparison between the PSE values for the Blur and Cornsweet profile (when matched with the Block profile) showed that PSE values were significantly higher for the Blur profile compared to the Cornsweet profile. These findings are consistent with the previous psychophysics experiment conducted, in which we found that thresholds for a Detection and Segmentation Task were lower for the Block profile compared to the Cornsweet profile, and the Blur profile had the highest thresholds (i.e. worse performance).

The fact that the Blur profile is the least salient can be explained by studies that have found poorer performance when orientation change occurs over a larger distance. According to Nothdurft (1985), the structure gradient of a texture stimulus takes into account not only the orientation contrast ($\Delta\theta$), but also the distance over which the change in orientation occurs (Δx). If $\Delta\theta$ is held constant, performance will decline while Δx increases. Therefore, a stimulus like the Blur profile that has orientation change occurring over a larger distance would not have a clearly defined edge boundary, which results in the saliency of the stimulus being reduced. However, stimuli like the Block and Cornsweet profiles that have a rapid change in orientation over a short distance would produce a clearly defined edge boundary,

which would have high saliency. The Block profile is still the most salient because it has orientation contrast at **both** figure edge and centre, instead of just at the figure edge like the Cornsweet profile. This pattern of results (Block > Cornsweet > Blur) is similar to what we have previously found in the psychophysical experiments, suggesting a commonality between the different tasks.

Finally, Deubel and Frank (1991) found that saccades are generated only after the figure has been segregated from the ground. To add to this, when we modelled the results using each participant's individual PSE values, we found that the size of the IR is similar to what we have found previously with the psychophysical experiments. These findings suggest that once the mechanism segregates the figure-ground textures, the saccadic system is then able to access the information produced by this mechanism to produce a saccadic response.

Chapter 11

General Discussion

In this thesis, we have used psychophysical (Chapters 4 – 6) and eye tracking techniques (Chapters 8 – 10) to investigate the role of texture edges in orientation-based figure-ground segregation. The consensus from all the experiments presented is that figure-ground segregation is not solely reliant on extraction of texture edges, but also regions beyond the immediate vicinity of the edge. This will be discussed further throughout this chapter, where we will first discuss the overall findings of the psychophysical experiments (Section 11.1), followed by the eye tracking experiments (Section 11.2). Subsequently, we will discuss the limitations of the work presented in this thesis and also further research that can be done in line with the research already conducted. We will finish with a concluding remark about the outcome of the collective research done.

11.1 Psychophysics Studies

Although many studies have investigated texture segregation using orientation-based stimuli, these figure-ground textures typically have an abutting edge i.e. an abrupt change in texture property to signal the figure border. Additionally, there are several studies that have shown texture segregation is **still** possible with smooth variations in texture i.e. no abutting edge (e.g. Gurnsey & Laundry 1992; Kingdom, Keeble, & Moulden, 1995; Kingdom & Keeble, 1996; Ben-Shahar & Zucker, 2004). The question we put forth here is whether there is a mechanism that can explain the results of both these types of stimuli? Thus, the psychophysics experiments presented in this thesis (Chapters 4 – 6) were conducted with the

general aim of investigating whether the mechanisms that mediate texture based segregation are edge-based or region-based.

11.1.1 Edge-based or Region-based Mechanisms?

A common feature of the psychophysics experiments (and most of the eye tracking experiments) is the type of stimuli we used. We constructed figure-ground texture patterns that had different texture variations i.e. the orientation profiles. The Block profile exemplified a stimulus with an abrupt texture variation, while the Blur profile exemplified a stimulus with a smooth texture variation. We also constructed a Cornsweet profile that had an abrupt texture change at the edge, which tapers off further away from the edge i.e. the orientation of the elements was the same within the centre of the figure and the ground region.

Briefly, an edge-based mechanism detects rapid discontinuities within a texture image, while a region-based mechanism groups together neighbouring elements that are similar (Wolfson & Landy, 1998). If texture segregation is mediated by an edge-based mechanism, it would be expected that performance for the Block and Cornsweet profiles will be the same owing to both profiles having orientation contrast information at the edge of the figure. Additionally, we would expect that performance for the Blur profile would be severely impaired (or perhaps not even possible) if only an edge-based mechanism was employed. On the other hand, if texture segregation is mediated by region-based mechanism, we would expect that segregation of the Blur profile will be possible.

Another aspect to take into account is the filling-in mechanism (e.g. Caelli, 1985; Caputo, 1998; Lamme et al., 1999; Scholte et al., 2008). For example, Caputo (1998) investigated the time course of texture filling-in using a masking paradigm. The findings were that the orientation defined line elements within the masking square appeared less bright because the

mask prevented filling-in from occurring. The effects of the size of the segregating texture and the ISI between the texture and the mask were investigated. The results were interpreted as demonstrating that texture filling-in started from the border (when orientation contrast was sufficiently high), and flowed inwards to the centre of the figure.

Lamme et al., (1999) also observed this sequence of edge detection followed by region filling-in. V1 neuronal activity in awake macaque monkeys was measured when receptive fields fell inside and outside of a figure. They found evidence of contextual modulation i.e. stronger signal strength to figure than ground region, which begins at ~70ms, but only for receptive fields near the edge border. By ~115ms, responses to the border have peaked, with only moderate responses to the figure centre. Peak neuronal responses to both figure edge and centre were observed at ~ 150ms. The authors concluded that the temporal sequence observed were in support of a filling-in mechanism of texture segregation. This finding of boundary detection preceding surface filling-in has been demonstrated by several other studies (Caputo & Casco, 1999; Scholte et al., 2008; Poort et al., 2012).

Therefore, we now have three possibilities: 1) edge-based mechanism, 2) region-based mechanism, and 3) edge-based and filling-in mechanism. In this context, an edge-based mechanism would operate over the immediate border of the figure edge to extract information. Moreover, a distinction can be made between a region-based mechanism and a filling-in mechanism, where region-based mechanisms instantly group together elements based on some similarity within the region, and filling-in mechanisms spread across a region slowly over time.

Thus, if filling-in occurs after edge extraction (or if a border edge is absent, filling-in occurs throughout the entire texture stimulus), performance for the Block profile should be better than the Cornsweet profile at longer stimulus durations, owing to the additional information

present at the centre of the figure for the Block profile. However at short stimulus durations, when region filling-in has **not** occurred, performance for the Block and Cornsweet profile should be the same. In sum, a combination of an edge-based and filling-in mechanism should yield results where performance for the Block and Cornsweet profiles are similar at short durations, but the Block profile will be better than the Cornsweet profile at long stimulus durations.

A consistent finding across Experiments 1 – 6 was that performance for the Block profile was almost always better than the Cornsweet and Blur profiles. Thus, an edge-based mechanism that extracts information from the elements immediately adjacent to the figure border cannot fully explain the results. Experiment 1 also showed that the Block profile constantly outperformed the Cornsweet profile at all stimulus durations tested. Thus, we find that 1) texture segregation is not mediated solely by an edge-based mechanism, and 2) there is no evidence of a filling-in mechanism. This leaves one other possibility – a region-based mechanism. However, if we were to assume that this region-based mechanism operates in a manner where regions that are similar to each other will be grouped together (Wolfson & Landy, 1998), we should have found that performance for the Block profile is similar to the Blur profile. A more plausible explanation is the presence of a region-based mechanism that integrates information over a large area of the visual field. This will be discussed in the following section.

11.1.2 Implications for Second-Order Texture Filters

As we discussed in the above section, neither an edge-based mechanism, nor an edge-based mechanism in conjunction with a filling-in mechanism could fully explain the results. We thus posit the presence of a mechanism that can integrate information over a large area of the visual field. We attempt to model this 'large area' by assuming that a common mechanism is

used to segregate textures with both abrupt and smooth texture variation (Kingdom & Keeble, 1996). We refer to this as the integration region (IR). The IR models the second filtering stage of the filter-rectify-filter model of texture segregation. That is, large-scale (i.e. large receptive fields) texture filters are used to detect orientation discontinuities.

Experiment 1 showed that the size of the IR did not increase with longer stimulus durations. We interpreted this result as evidence of the second-order filter having a fixed sized, with no indication once again of a filling-in mechanism. We tested this in Experiment 2, where we varied the parameter that determines the curve of the slope that varies the orientation change of the Blur and Cornsweet profiles. Assuming the second-order filter has a fixed size, the information of orientation change within a specific region size would differ according to the various profiles. For example, if the size of the filter is fixed at 3 raster widths, a steeper slope of the Blur profile would result in a greater average orientation contrast between the 3 rows on either side of the texture border. However, a shallow slope i.e. orientation change occurs over a further distance, would produce a lower average orientation contrast between the 3 rows on either side of the texture border. The threshold results of Experiment 2 were in line with what we would expect if a fixed size second-order texture filter mediates texture segregation.

In the texture segregation literature, there is evidence of multiple second-order texture filters of different sizes tuned to the luminance spatial frequency of the first-stage inputs (e.g. Sutter, Sperling, & Chubb, 1995; Kingdom & Keeble, 1999). For example, Kingdom and Keeble (1999) found high luminance spatial frequency would result in a small second-stage filter, while low luminance spatial frequency would result in a bigger second-stage filter. We therefore ran a series of experiments to investigate what properties of a stimulus would also affect the size of these second-stage filters (thus the IR). We manipulated the size of the

figure (Experiment 3), the figure aspect ratio (Experiment 4), the age group of the participants and the luminance spatial frequency (Experiment 5), and also the viewing distance (Experiment 6).

IR results for those experiments were that the size of the second-order filter, as measured in object units not retinal units, remained the same with most of those manipulations. The exception for that was the luminance spatial frequency, where we found larger IRs with higher luminance spatial frequencies. This is in contrast to the findings of Kingdom and Keeble (1999). The only explanation we have for this difference is the nature of the model itself. Even though both their model and ours are represented as physical units on the screen, their model was in terms of the texture spatial frequency that was most sensitive (TSF_{min}) while ours was the integration size in which the thresholds of the three different orientation profiles used were most similar. Thus, it is conceivable that our results do not agree based on this difference.

11.2 Eye Tracking Studies

Psychophysical studies that investigate orientation-based figure-ground segregation typically limit the presentation time of the stimulus for the purpose of reducing the influence of eye movements during the task. However, He and Kowler (1992) showed that participants performed better in a texture discrimination task when they were allowed to make eye movements compared to when they had to maintain fixation at only one point. Thus, preventing participants from making eye movements, as we did in the psychophysics experiments, may well be limiting our understanding of the segregation process. In the eye tracking experiments presented in this thesis (Chapters 8 – 10), the general aim was to investigate the way information from a texture target is processed in order to provide a signal for eye movement control.

11.2.1 Saccade Programming

The essential feature common to the experiments presented in Chapter 9 (Experiments 8 – 10) is that saccadic eye movements were to be directed to targets embedded in a background structure. The aim of those studies was to determine whether the reference position of the saccade was computed by pooling information across the whole target shape, or by selecting a salient location within the target based orientation contrast cues. In the first experiment of that chapter (Experiment 8), we manipulated the orientation profiles of the target figure, whereby each profile had cues of orientation contrast at different parts of the figure. Despite this, we found that the orientation profiles did not influence the saccade landing position, which was on average 85% relative to figure centre. This result suggests that saccades are guided by the representation of the entire target figure, **not** by any particular salient location within the target.

That still raises the question of how saccade destinations are computed for targets embedded in a background structure. Could it be based on the spatial pooling i.e. averaging of the target endpoints as suggested by Findlay (1982), Deubel et al., (1988), and Ploner et al., (2004), or perhaps based on the whole figure shape as suggested by Melcher and Kowler (1999)? In Experiment 10, we addressed this question by using triangle-shaped target figures. Unlike the rectangle-shaped figures used in Experiment 8, the centre (midpoint of the target width) and centre of gravity (COG, average position of all points in the figure) of the triangle-shaped targets do not coincide with each other. Thus, if saccades were directed to the centre, it would imply that the destination of the saccade was computed based on the spatial pooling of just the two endpoints of the target width. Alternatively, if saccades were directed to the COG, it would imply that the saccades were programmed based on spatial pooling of information across the whole target figure.

The results of Experiment 10 showed that saccades are directed to the centre of gravity (COG), which supports the notion that saccades are programmed based on the representation of the target figure as a whole object. This is in accordance with the findings of Melcher and Kowler (1999). However, they specifically instructed their participants to "look at the target as a whole, rather than aim the saccade to a particular place within" (p. 2933) and the targets in their study were presented in an impoverished visual environment (i.e. no background). It is abundantly clear that the natural visual environment normally constitutes complex targets in which the foreground "figure" has to first be segregated from the background structure. Our findings therefore extend theirs by claiming: 1) even when participants were not specifically told to look at the target as a whole, that is what the participants did, and 2) even when a target is embedded in a background, saccade destination is computed based on the representation of the entire shape of the texture figure by first segregating the figure from the background.

In fact, in Experiment 11, we used participants' PSE values, which were the degrees of orientation contrast needed to equate the saliency across the three orientation profiles, to compute the IR. We found that the IR size for that experiment was consistent with the size of the IR for the comparable psychophysics experiment. We therefore propose that early mechanisms that segregate figure-ground textures provide input to areas that guide oculomotor planning. This is supported by findings that show early visual areas of the cortex (V1, V2, V3, and the middle temporal area) project to the superficial layers of the superior colliculus (Cerkevich, Lyon, Balaram, & Kaas, 2014), and also by findings that show neuronal activity in the V1 provides internal evidence of a stimulus that guides saccades to figure-ground texture stimuli (Supèr & Lamme, 2007).

Saccade amplitudes of the experiments presented in Chapter 9 showed that the saccadic system is adept at accurately planning the destination of a saccade so that it lands on the target, albeit with a certain degree of undershoot respective to the target centre. Saccade metrics also showed that saccade destinations are more accurate **and** precise for smaller targets compared to larger targets (Experiment 8 vs. 9). Similar effects have been observed when participants made saccades to solitary targets (Kowler & Blaser, 1995; Ploner, Ostendorf, & Dick, 2004). Even in double-target experiments, where the saccade landing position is to an intermediate position between the two targets, as the distance between the two target increases, there is a greater undershoot to the midpoint of the two targets (Findlay, 1982). Since the spatial pooling of visual information has been linked with the saccadic system computing a reference location for the saccade destination (e.g. Findlay, 1982; Kowler & Blaser, 1995; Ploner, Ostendorf, & Dick, 2004), a decline in accuracy and precision of saccades are most likely the result of the more complex processing required to spatially pool information across a larger region.

Indeed, we find that saccades are more accurate for nearer targets than further targets (evident from Experiments 9 & 10, but not 8). Therefore, it suggests that spatially pooling information from a target that is further away from fixation is also more complex. But, if pooling of information for far targets is more complex, why do we find that saccades are more precise for targets further away? Inspection of graphs that show saccade amplitudes plotted as frequency histograms (see Figures 9.3, 9.12 and 9.18 in Chapter 9), reveal that for targets further from fixation, there are fewer saccades to the border edge further from fixation. For targets closer to fixation, the spread of saccades is throughout the entire figure region. That is, further targets are not more precise relative to the figure centre, but precision is merely an artifact of fewer saccades (or none at all) landing on the further regions of the figure.

Furthermore, we showed that saccade latencies were faster for smaller targets (262ms) than larger targets (293ms). Similar effects have been observed by Ploner et al., (2004), but not Kowler and Blaser (1995). As suggested by Ploner and colleagues, larger targets require more complex spatial pooling of visual information, which results in greater saccade latencies to the target. Unlike our experiments, Kowler and Blaser specifically instructed participants to increase saccade latencies in favour of greater saccade accuracy, which could explain why they observe no difference in saccade latencies for the different target sizes.

Finally, we ask why were participants able to localise the target figures in Experiments 8 – 10, but not Experiment 7? In Experiment 7, not only did we use figure-ground stimuli with low orientation contrast (i.e. less detectable), we also included stimuli where there was no figure embedded in the background (catch trials). As such, when we asked participants to locate the figure patch, they would have adopted a completely different strategy. In fact, when we analysed the location of first fixation (i.e. saccade destination), we found that the saccade amplitudes were on average 84.5% relative to the centre of the texture grid. It implies that first fixations were not directed to the target figure, but to the centre of the texture grid with approximately 15% undershoot.

To sum up, our experiments have demonstrated that: 1) our visual system is able to reliably segregate a texture figure embedded in a background, and accurately plan a saccade to the figure, 2) saccade destination is computed based on the representation of the entire shape of the texture figure, not just the salient cues, 3) saccade destination is computed by spatially pooling information of the target, 4) increased size and distance of the target increases the complexity of spatial pooling, which affects the accuracy and latency of the saccade, and 5) areas that guide oculomotor planning receive input from early visual mechanisms that segregate figure-ground textures.

11.2.2 Effects of Target Saliency

In the experiments presented in Chapter 8 – 10, we have observed that some characteristics of eye movements seem to be influenced by the saliency of the target. Saliency here refers to the quality of being particularly detectable. By manipulating the orientation contrast of the figure-ground texture stimuli, we can alter the detectability of a texture figure. This is evident by studies that show figures segregate more easily from the background when orientation contrast is high (e.g. Nothdurft, 1985b).

In Experiment 11, we presented participants with two figure-ground texture grids on either side of a central fixation point. Participants were instructed to move their eyes to the texture grid that was more noticeable. By measuring the PSE, we were able to measure the degree of orientation contrast needed by all three orientation profiles to have the same level of saliency. We showed that the degree of orientation contrast needed to match the Block profile was higher by 32% for the Cornsweet profile and 121% for the Blur profile. We therefore claim that Block profile, which has orientation contrast information at both the figure edge and centre, is the most salient orientation profile.

In Experiment 9, we measured the saccade latencies of participants to the various orientation profiles. We found that the saccade latencies to the Blur profile were the longest compared to the Block and Cornsweet profiles. Supèr, Spekreijse, and Lamme (2003) showed that stronger contextual modulation (stronger neuronal response to the figure region compared to the ground region) would result in faster saccade latencies. This is a result of a strong signal of modulation producing a rapid build-up in the motor system (prior to saccade onset), which results in shorter saccade latencies (Supèr & Lamme, 2007). Thus, when a target is less salient (i.e. the Blur profile), the time taken to initiate the saccade is longer. A similar effect was observed by Nothdurft and Parlitz (1993), who showed that orientation defined targets

did not produce express saccades due to the signal strength being weaker than luminance defined targets

If we claim that less salient targets will produce longer saccade latencies, why do we observe no difference of saccade latencies in Experiment 7? The simple explanation is that the first saccade to the stimuli with that specific experimental setup and task demands is to a point that was unrelated to the target figure, but was in relation to the centre of the entre texture grid. That is, the saliency of the figure was not driving the saccade destination, but the entire texture grid. As such, saccade latencies would not have been influenced whether the figure was salient (Block profile, or high orientation contrast stimuli) or not.

Other metrics of eye movements – such as fixation count and summed fixation duration, also showed that when target saliency increases, the number of fixations and summed fixation durations decrease (Experiment 7). When a target is more salient i.e. detectable, the figure segregates more readily and participants are able to localise the target figure with greater ease. As a consequence, they require fewer and shorter fixations before responding. Similar effects have been found by Van Humbeeck et al., (2013) with a contour integration task. They found that the number of fixations and total viewing time increased when the contour target was less salient.

11.2.3 Which Part of a Figure is Extracted?

In the experiments presented in Chapter 8, we investigated how eye movements were influenced by figure-ground texture stimuli. We were particularly interested in circumstances where the figure target was easy and difficult to detect. The difficulty of detecting the target was manipulated by varying the orientation contrast between figure and ground, where figures segregate more easily with high orientation contrast. We analysed participants'

fixation counts and summed fixation durations to the centre and edge region of the figure. We area-normalized both measures by dividing the proportion responses to the region by the proportional size of the region. Therefore, if area-normalized responses are greater than 1, it indicates that the responses to the region are occurring at a rate higher than would be expected by chance (e.g. Fletcher-Watson et al., 2008; Bindemann et al., 2009).

The key finding of the experiments that measure area-normalized fixation count and summed fixation duration is that the centre region is looked at more often and for longer durations compared to the edge region. This is irrespective of the orientation profiles of the target figure (pilot study, Experiment 7 & 8), or the task to locate the centre/edge of the figure (Experiment 7). Therefore, eye movements are predominantly directed to the centre region of the figure, regardless of the cues of orientation contrast being present or absent in that region. However, area-normalized scores were still greater than 1 to the edge region, suggesting that even though it is to a lesser extent than the centre region, the edge region is still fixated on at a rate higher than expected by chance.

In all, we find that texture stimuli that have orientation contrast information at **both** the edge and centre of the figure are most salient in driving eye movements. Nevertheless, there is a bias to **looking** more frequently and for longer durations to the centre region of a texture target, irrespective of the orientation contrast information or task demand.

11.3 Limitations and Future Work

The experiments presented in this thesis are not without limitations. For the psychophysics studies especially, one major limitation is the difficulty of the tasks itself. For the experienced observers, this was not much of an issue, but it was for the naïve observers despite the extensive practice sessions they undertook. Moreover, since the measure used was of

participants' orientation contrast thresholds, there was an upper limit to an acceptable threshold (90°), and anything higher would be excluded. Consequently, we had results where a number of data points were missing from the participants, especially the naïve observers' data, making interpretations of the results slightly weaker.

In fact, in Experiment 5 (Section 6.2), 7 out of the 15 older participants were unable to produce physically meaning thresholds for the Segmentation task. This is in contrast to only 1 younger participant that was unable to produce physically meaningful thresholds. Thus, even though we found no group difference in thresholds of the young and old participants, our sample size was rather unequal. All things considered, even though the task difficulty seems to have reduced our data sample, this mostly affected the Segmentation task. In most instances, we have been able to get physically meaningful threshold for the Detection task, as it is a simpler task (as observed by the lower thresholds).

For the eye tracking experiments, only after conducting two studies (Chapter 8) did we realize that the experimental design does strongly influence the eye movements participants made. In Chapter 9, we modified our experimental design to suitably investigate our research question. However, this tells us that the type of task, and possibly also how participants are instructed, could play a role in the eye movement patterns. As such, we should be cautious when comparing results in which the experimental design is different.

Finally, consideration has to be given to the type of stimuli used in our experiment. As mentioned in Chapter 1, a single visual feature practically never defines the natural visual world. However, the stimuli used here were of a single feature – orientation. We have made the choice to study texture properties in isolation of single features, while other studies (e.g. Itti & Koch, 2001; Parkhurst, Law, & Neibur, 2002) have also used natural scenes to investigate what low-level feature is extracted from a scene. As a result, the stimuli here lack

a certain degree of ecological validity. However, the figure-ground stimuli, though nothing like what we see in the real world, constituted of angular figure shapes, which are what we commonly see in the urban environment (e.g. building, doors, house).

The work presented here still leaves much scope for further investigations in psychophysics and eye tracking studies into orientation-based texture segregation. For the psychophysics studies, we have conducted a series of experiments (Chapters 4 – 6) to determine what would affect the size of the IR. We have found that IR changes for the different tasks (segmentation and detection) performed, and also the luminance spatial frequency of the Gabor patches. However, other factors such as stimulus duration, orientation jitter of line elements, positional arrangement of line elements, figure sizes, figure aspect ratios, age group and viewing distance do not affect the IR. To extend to this, we can vary the line spacing (i.e. density, which is the number of elements per unit area) of the stimuli to determine if the IR changes based on visual angle or the number of rows. Moreover, in Experiment 3, where we varied the size of the texture figure, we found that the overall size of the IR was larger compared to the other experiments. We attributed this to the texture grid being larger at 40×40 elements, instead of the usual 32×32 elements. In a follow-up experiment, this variation could explicitly be tested.

Additionally, our results of Experiments 6 & 7, where we found that Gabor patches with higher luminance spatial frequency had larger integration regions, is in contrast to what Kingdom and Keeble (1999) found. We theorised that the difference in results could be due to the different mechanisms used in the two experiments. A simple way to test this would be to conduct an experiment using the three orientation profiles, but instead of the figure detection or segmentation task, we would measure the orientation modulation function (OMF). That is, we would measure the sensitivities of the orientation modulation as a

function of different texture spatial frequencies (number of cycles of orientation modulation per screen).

It would also be interesting to extend our research by investigating other types of figure-ground texture i.e. not just orientation-based figure-ground texture. That is, instead of a figure that is defined by a difference in orientation from the background, it could be defined by a difference in spatial frequency, contrast, or even direction of motion.

In the eye tracking experiments presented in this thesis, we established that saccades are directed to the COG of the target shape. However, Melcher and Kowler (1999) claim that saccades landed on the centre of area (COA, which is the centre of mass assuming that the shape has a uniform density) of a target, not the COG (average position of all elements that make up the target). Our experiments were not able to test this distinction as we used figure-ground texture with uniform densities. A follow-up experiment could investigate this by using targets with line elements that varied in density across the figure. If saccades land closer to the point where there are more elements (high density), it would imply that saccades to texture targets are computed based on the COG, not COA. Moreover, as mentioned above, our figure-ground stimuli comprised figure shapes that are rather common. It would be interesting to see how eye movements are directed to irregular and uncommon shapes (e.g. fractals, blobs)

Finally, it would be especially interesting to use single-cell recordings in primates to investigate neuronal activity in V1 using the Blur and Cornsweet profiles. Lamme's work (e.g. Lamme, 1995; Lamme et al., 1999) has used figure-ground stimuli with an abrupt texture variation (comparable to the Block profile). The results showed that neuronal activity was higher when the receptive field was on the figure region compared to the ground region. With our Cornsweet profile – where there is an abrupt texture edge, but orientation of the line

elements in the centre of the figure is the same as the ground – would neuronal activity still occur? And if so, would it only occur if the receptive field was on the edge of the figure? Likewise, with the Blur profile, would neuronal activity occur at the edge of the figure (where the orientation change is minimal), or only the centre? And would the signal strength be as strong as the Block profile?

11.4 Conclusion

In summary, the work presented in this thesis adds to the growing literature on the mechanisms of orientation-based texture segregation. It confirms the role of a region-based mechanism that is involved in the segregation process. Specifically, our results are best explained by the presence of a large-scale second-order texture filter that mediates texture segregation. We also propose that the output of this filter feeds into other areas of the brain that then guides saccades to texture targets. These saccades land accurately on the target based on the representation of the whole shape of the figure.

Reference

Allen, H. A., Hutchinson, C. V., Ledgeway, T., & Gayle, P. (2010). The role of contrast sensitivity in global motion processing deficits in the elderly. *Journal of Vision*, *10*(10), 1–15. https://doi.org/10.1167/10.10.15

Appelbaum, L. G., Ales, J. M., & Norcia, A. M. (2012). The time course of segmentation and cue-selectivity in the human visual cortex. *PLoS ONE*, *7*(3), 1–12. https://doi.org/10.1371/journal.pone.0034205

Arden, G. B. (1978). The importance of measuring contrast sensitivity in cases of visual disturbance. *British Journal of Ophthalmology*, *62*(4), 198–209.

Attar, C. H., Hamburger, K., Rosenholtz, R., Gotzl, H., & Spillmann, L. (2007). Uniform versus random orientation in fading and filling-in. *Vision Research*, *47*(24), 3041–3051. https://doi.org/10.1016/j.visres.2007.07.022

Beck, J. (1966). Effect of orientation and of shape similarity on perceptual grouping. *Perception & Psychophysics*, *1*(9), 300–302. https://doi.org/10.3758/BF03215792

Beck, J. (1982). Textural segmentation. In J. Beck (Ed.), *Organization and Representation in Perception* (pp. 285–317). Hillsdale, NJ: Erlbaum.

Beck, J., Sutter, A., & Ivry, R. (1987). Spatial frequency channels and perceptual grouping in texture segregation. *Computer Vision, Graphics, and Image Processing*, *37*(2), 299–325. https://doi.org/10.1016/S0734-189X(87)80006-3

Bell, J., Gheorghiu, E., Hess, R. F., & Kingdom, F. A. (2011). Global shape processing involves a hierarchy of integration stages. *Vision Research*, *51*(15), 1760-1766.

Ben-Shahar, O. (2006). Visual saliency and texture segregation without feature gradient. *PNAS*, *103*(45), 15704–15709.

Ben-Shahar, O., & Zucker, S. W. (2004). Sensitivity to curvatures in orientation-based texture segmentation. *Vision Research*, *44*(3), 257–277. https://doi.org/10.1016/j.visres.2003.08.018

Ben-Yosef, G., & Ben-Shahar, O. (2008). Curvature-based perceptual singularities and texture saliency with early vision mechanisms. *JOSA A*, *25*(8), 1974–1993.

Bergen, J. R. (1991). Theories of visual texture perception. In D. Regan (Ed.), *Vision and visual dysfunction* (pp. 114–134). New York: Macmillian.

Bergen, J. R., & Adelson, E. H. (1986). Visual texture segmentation based on energy measures. *J. Opt. Soc. Am.*, *A*(3), 98–99.

Bergen, J. R., & Adelson, E. H. (1988). Early vision and texture perception. *Nature*, *333*(6171), 363–364.

Bergen, J. R., & Landy, M. S. (1991). Computational modeling of visual texture segregation. In M. S. Landy & J. A. Morshon (Eds.), *Computational models of visual processing* (pp. 253–271). Cambridge, MA: MIT Press.

Bertone, A., Guy, J., & Faubert, J. (2011). Assessing spatial perception in aging using an adapted Landolt-C technique. *NeuroReport*, *22*(18), 951–955. https://doi.org/10.1097/WNR.0b013e32834d2f49

Betts, L. R., Taylor, C. P., Sekuler, A. B., & Bennett, P. J. (2005). Aging reduces center-surround antagonism in visual motion processing. *Neuron*, *45*(3), 361–366. https://doi.org/10.1016/j.neuron.2004.12.041

Bindemann, M., Scheepers, C., & Burton, A. M. (2009). Viewpoint and center of gravity affect eye movements to human faces. *Journal of Vision*, *9*(2), 1–16. https://doi.org/10.1167/9.2.7

Blakemore, C., & Campbell, F. W. (1969). On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images. *The Journal of Physiology*, *203*(1), 237–260. https://doi.org/10.1113/jphysiol.1969.sp008862

Caelli, T. (1985). Three processing characteristics of visual texture segmentation. *Spatial Vision*, *1*(1), 19–30. https://doi.org/10.1163/156856885X00044

Caelli, T., & Julesz, B. (1978). On perceptual analyzers underlying visual texture discrimination: Part I. *Biological Cybernetics*, *28*(3), 167–175. https://doi.org/10.1007/BF00337138

Campbell, F. W., & Robson, J. G. (1968). Application of Fourier analysis to the visibility of gratings. *The Journal of physiology*, *197*(3), 551.

Caputo, G. (1998). Texture brightness filling-in. *Vision Research*, *38*(6), 841–851. https://doi.org/10.1016/S0042-6989(97)00237-X

Caputo, G., & Casco, C. (1999). A visual evoked potential correlate of global figure-ground segmentation. *Vision Research*, *39*(9), 1597–1610. https://doi.org/10.1016/S0042-6989(98)00270-3

Carbal, B., & Leedom, L. (1993). Imaging vector fields using line integral convolution. In *Proceedings of the 20th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH 1993* (pp. 263–270). New York: ACM Press. https://doi.org/10.1145/166117.166151

Carpenter, R. H. S. (1981). Oculomotor procrastination. In D. Fisher, R. Monty, & J. Senders (Eds.), *Eye movements: Cognition and Visual Perception* (pp. 237–246). Hillsdale, NJ: Lawrence Erlbaum.

Carpenter, R. H. S., & Williams, M. L. L. (1995). Neural computation of log likelihood in control of saccadic eye movements. *Nature*, *377*(6544), 59–62. https://doi.org/10.1038/377059a0

Casco, C., Barollo, M., Contemori, G., & Battaglini, L. (2017). The effects of aging on orientation discrimination. *Frontiers in Aging Neuroscience*, *9*, 45. https://doi.org/10.3389/fnagi.2017.00045

Casco, C., Grieco, A., Campana, G., Corvino, M. P., & Caputo, G. (2005). Attention modulates psychophysical and electrophysiological response to visual texture segmentation in humans. *Vision Research*, *45*, 2384–2396. https://doi.org/10.1016/j.visres.2005.02.022

Casco, C., Robol, V., Barollo, M., & Cansino, S. (2011). Effects of aging on visual contour integration and segmentation. *Investigative Ophthalmology and Visual Science*, *52*(7), 3955–3961. https://doi.org/10.1167/iovs.10-5439

Castelhano, M. S., Mack, M. L., & Henderson, J. M. (2009). Viewing task influences eye movement control during active scene perception. *Journal of Vision*, *9*(3), 1–15. https://doi.org/10.1167/9.3.6

Cerkevich, C. M., Lyon, D. C., Balaram, P., & Kaas, J. H. (2014). Distribution of cortical neurons projecting to the superior colliculus in macaque monkeys. *Eye and brain*, *6*(Suppl 1), 121.

Chennamma, H. R., & Yuan, X. (2013). A Survey On Eye-Gaze Tracking Techniques. *Indian Journal of Computer Science and Engineering (IJCSE)*, *4*(5), 388–393.

Coren, S., & Hoenig, P. (1972). Effect of non-target stimuli upon length of voluntary saccades. *Perceptual and Motor Skills*, *34*(2), 499–508. https://doi.org/10.2466/pms.1972.34.2.499

Cornsweet, T. (1970). *Visual Perception*. New York: Academic Press.

Crane, H. D. (1994). The Purkinje Image Eyetracker, Image Stabilization, and Related Forms of Stimulus Manipulation. In D. H. Kelly (Ed.), *Visual Science and Engineering Models and Applications* (pp. 15–89). New York: Mercel Dekker.

Delahunt, P. B., Hardy, J. L., & Werner, J. S. (2008). The effect of senescence on orientation discrimination and mechanism tuning. *Journal of Vision*, *8*(3), 1–9. https://doi.org/10.1167/8.3.5

Del Viva, M. M., & Agostini, R. (2007). Visual spatial integration in the elderly. *Investigative Ophthalmology and Visual Science*, *48*(6), 2940–2946. https://doi.org/10.1167/iovs.06-0729

Deubel, H., Findlay, J., Jacobs, A., & Brogan, D. (1988). Saccadic eye movements to targets defined by structure differences. In G. Lüer, U. Lass, & J. Schallo-Hoffmann (Eds.), *Eye movement research: Physiological and psychological aspects* (pp. 107–145). Toronto: Hogrefe.

Deubel, H., & Frank, H. (1991). The latency of saccadic eye movements to texture-defined stimuli. In R. Schmid & D. Zambarbieri (Eds.), *Oculomotor control and cognitive processes* (pp. 369–384). North-Holland: Elsevier.

DeValois, R. L., & DeValois, K. K. (1990). *Spatial Vision*. New York: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195066579.001.0001

Dodge, R., & Cline, T. S. (1901). The angle velocity of eye movements. *Psychological Review*, *8*(2), 145–157. https://doi.org/10.1037/h0076100

Duchowski, A. (2007). *Eye tracking methodology: Theory and practice* (2nd ed.). London: Springer-Verlag. https://doi.org/10.1007/978-1-84628-609-4

Faubert, J. (2002). Visual perception and aging. *Canadian Journal of Experimental Psychology*, *56*(3), 164–176. https://doi.org/10.1037/h0087394

Field, D. J., Hayes, A., & Hess, R. F. (1993). Contour integration by the human visual system: Evidence for a local "association field." *Vision Research*, *33*(2), 173–193. https://doi.org/10.1016/0042-6989(93)90156-Q

Findlay, J. M. (1982). Global visual processing for saccadic eye movements. *Vision Research*, *22*(8), 1033–1045. https://doi.org/10.1016/0042-6989(82)90040-2

Fischer, B., & Boch, R. (1983). Saccadic eye movements after extremely short reaction times in the monkey. *Brain Research*, *260*(1), 21–26. https://doi.org/10.1016/0006-8993(83)90760-6

Fletcher-Watson, S., Findlay, J. M., Leekam, S. R., & Benson, V. (2008). Rapid detection of person information in a naturalistic scene. *Perception*, *37*(4), 571–583. https://doi.org/10.1068/p5705

Gabor, D. (1946). Theory of communication. Part 1: The analysis of information. *Journal of the Institution of Electrical Engineers-Part III: Radio and Communication Engineering*, *93*(26), 429–441. https://doi.org/10.1049/ji-1.1947.0015

Georgeson, M. A. (1973). Spatial frequency selectivity of a visual tilt illusion. *Nature*, *245*(5419), 43–45. https://doi.org/10.1038/245043a0

Gescheider, G. A. (1997). *Psychophysics: The Fundamentals* (3rd ed.). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Gilchrist, I. D. (2011). Saccades. In S. Liversedge, I. Gilchrist, & S. Everling (Eds.), *The Oxford Handbook of Eye Movements* (pp. 85–94). New York: Oxford University Press.

Giora, E., & Casco, C. (2007). Region- and edge-based configurational effects in texture segmentation. *Vision Research*, *47*, 79–886. https://doi.org/10.1016/j.visres.2007.01.009

Govenlock, S. W., Taylor, C. P., Sekuler, A. B., & Bennett, P. J. (2009). The effect of aging on the orientational selectivity of the human visual system. *Vision Research*, *49*(1), 164–172. https://doi.org/10.1016/j.visres.2008.10.004

Gupta, S., & Sproull, R. F. (1981). Filtering Edges for Gray-Scale Displays. *Computer Graphics*, *15*(3), 1–5.

Gurnsey, R., & Laundry, D. (1992). Texture Discrimination With and Without Abrupt Texture Gradients. *Canadian Journal of Psychology*, *46*(2), 306–332.

Habak, C., & Faubert, J. (2000). Larger effect of aging on the perception of higher-order stimuli. *Vision Research*, *40*(8), 943–950. https://doi.org/10.1016/S0042-6989(99)00235-7

Hadad, B. S. (2012). Sensitivity of spatial integration to perceptual cues is preserved in healthy aging. *Vision Research*, *60*, 1–6. https://doi.org/10.1016/j.visres.2012.03.002

Hadad, B. S., & Kimchi, R. (2008). Time course of grouping of shape by perceptual closure: Effects of spatial proximity and collinearity. *Perception and Psychophysics*, *70*(5), 818–827. https://doi.org/10.3758/PP.70.5.818

Harrison, S. J., & Keeble, D. R. T. (2008). Within-texture collinearity improves human texture segmentation. *Vision Research*, *48*(19), 1955–1964. https://doi.org/10.1016/j.visres.2008.06.008

He, P., & Kowler, E. (1992). The role of saccades in the perception of texture patterns. *Vision Research*, *32*(11), 2151–2163. https://doi.org/10.1016/0042-6989(92)90076-U

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. Oxford: Oxford University Press.

Hubel, D. H., & Wiesel, T. N. (1977). Ferrier lecture: Functional architecture of macaque monkey visual cortex. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, *198*(1130), 1–59. https://doi.org/10.1098/rspb.1977.0085

Hunt, J. J., Mattingley, J. B., & Goodhill, G. J. (2012). Randomly oriented edge arrangements dominate naturalistic arrangements in binocular rivalry. *Vision Research*, *64*, 49–55. https://doi.org/10.1016/j.visres.2012.05.007

Hutchinson, C. V., Arena, A., Allen, H. A., & Ledgeway, T. (2012). Psychophysical correlates of global motion processing in the aging visual system: A critical review. *Neuroscience and Biobehavioral Reviews*, *36*(4), 1266–1272. https://doi.org/10.1016/j.neubiorev.2012.02.009

Hutchinson, C. V., Ledgeway, T., & Allen, H. A. (2014). The ups and downs of global motion perception: A paradoxical advantage for smaller stimuli in the aging visual system. *Frontiers in Aging Neuroscience*, *6*, 199. https://doi.org/10.3389/fnagi.2014.00199

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert hifts of visual attention. *Vision Research*, *40*(10–12), 1489–1506. https://doi.org/10.1016/S0042-6989(99)00163-7

Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature Reviews Neuroscience*, *2*(3), 194–203. https://doi.org/10.1038/35058500

Janik, S. W., Wellens, A. R., Goldberg, M. I., & Dell'Osso, J. F. (1978). Eyes as the center of focus in the visual examination of human faces. *Perceptual & Motor Skills*, *47*, 857–858.

Jingling, L., & Tseng, C. H. (2013). Collinearity impairs local element visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(1), 156–167. https://doi.org/10.1037/a0027325

Julesz, B. (1962). Visual Pattern Discrimination. *IEEE Transactions on Information Theory*, *8*(2), 84–92. https://doi.org/10.1109/TIT.1962.1057698

Julesz, B. (1981). Textons, the elements of texture perception, and their interactions. *Nature*, *290*, 91–97.

Kapoula, Z., & Robinson, D. A. (1986). Saccadic undershoot is not inevitable: Saccades can be accurate. *Vision Research*, *26*(5), 735–743. https://doi.org/10.1016/0042-6989(86)90087-8

Kastner, S., De Weerd, P., & Ungerleider, L. G. (2000). Texture Segregation in the Human Visual Cortex: A Functional MRI Study. *Journal of Neurophysiology*, *83*(4), 2453–2457.

Kingdom, F. A. A., Keeble, D., & Moulden, B. (1995). Sensitivity to orientation modulation in micropattern-based textures. *Vision Research*, *35*(1), 79–91. https://doi.org/10.1016/0042-6989(94)E0079-Z

Kingdom, F. A. A., & Keeble, D. R. T. (1996). A linear systems approach to the detection of both abrupt and smooth spatial variations in orientation-defined textures. *Vision Research*, *36*(3), 409–420. https://doi.org/10.1016/0042-6989(95)00123-9

Kingdom, F. A. A., & Keeble, D. R. T. (1999). On the mechanism for scale invariance in orientation-defined textures. *Vision Research*, *39*(8), 1477–1489. https://doi.org/10.1016/S0042-6989(98)00217-X

Kingdom, F. A. A., & Prins, N. (2010). *Psychophysics: A Practical Introduction*. San Diego, CA,: Elsevier Academic Press.

Klein, R. M. (2000). Inhibition of return. *Trends in Cognitive Neuroscience*, *4*(4), 138–147. https://doi.org/10.1016/S1364-6613(00)01452-2

Koch, C., & Ullman, S. (1985). Shifts in selective visual attention: Towards the underlying neural circuitry. *Human Neurobiology*, *4*(4), 219–227.

Kowler, E., & Blaser, E. (1995). The accuracy and precision of saccades to small and large targets. *Vision Research*, *35*(12), 1741–1754. https://doi.org/10.1016/0042-6989(94)00255-K

Kurylo, D. D. (2006). Effects of aging on perceptual organization: Efficacy of stimulus features. *Experimental Aging Research*, *32*(2), 137–152. https://doi.org/10.1080/03610730600553901

Lachapelle, J., McKerral, M., Jauffret, C., & Bach, M. (2008). Temporal resolution of orientation-defined texture segregation: A VEP study. *Documenta Ophthalmologia*, *117*, 155–162. https://doi.org/10.1007/s10633-008-9117-x

Lamme, V. A. F. (1995). The neurophysiology of figure-ground segregation in primary visual cortex. *The Journal of Neuroscience*, *15*(2), 1605–1615.

Lamme, V. A. F., Rodriguez-Rodriguez, V., & Spekreijse, H. (1999). Separate Processing Dynamics for Texture Elements, Boundaries and Surfaces in Primary Visual Cortex of the Macaque Monkey. *Cerebral Cortex*, *9*(4), 406–413. https://doi.org/10.1093/cercor/9.4.406

Lamme, V. A. F., Supèr, H., & Spekreijse, H. (1998). Feedforward, horizontal, and feedback processing in the visual cortex. *Current Opinion in Neurobiology*, *8*(4), 529–535. https://doi.org/10.1016/S0959-4388(98)80042-1

Land, M. F., & Hayhoe, M. M. (2001). In what ways do eyemovements contribute to everyday activities? *Vision Research*, *41*(25–26), 3559–3565.

Land, M., & Tatler, B. (2009). *Looking and Acting: Vision and eye movements in natural behaviour*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198570943.001.0001

Landy, M. S. (2013). Texture analysis and perception. In J. S. Werner & L. M. Chalupa (Eds.), *The new visual neurosciences* (pp. 639–652). Cambridge, Mass.: MIT Press.

Landy, M. S., & Bergen, J. R. (1991). Texture segregation and orientation gradient. *Vision Research*, *31*(4), 679–691. https://doi.org/10.1016/0042-6989(91)90009-T

Landy, M. S., & Graham, N. (2004). Visual Perception of Texture. In L. M. Chalupa & J. S. Werner (Eds.), *The Visual Neurosciences* (pp. 1106–1118). Cambridge, Mass: MIT Press.

Li, J., & Li, Z. (2008). Change detection is easier at texture border bars when they are parallel to the border: Evidence for V1 mechanisms of bottom-up salience. *Perception*, *37*, 197–206. https://doi.org/10.1068/p5829

Li, J., Tseng, C.-H., & Li, Z. (2013). Orientation is different: Interaction between contour integration and feature contrasts in visual search. *Journal of Vision*, *13*(3), 1–13.

Li, Z. (1999a). Contextual influences in V1 as a basis for pop out and asymmetry in visual search. *Proceedings of the National Academy of Sciences*, *96*(18), 10530–10535.

Li, Z. (1999b). Visual segmentation by contextual influences via intra-cortical interactions in the primary visual cortex. *Network: Computation in Neural Systems*, *10*(2), 187–212.

Li, Z. (2000). Pre-attentive segmentation in the primary visual cortex. *Spatial Vision*, *13*(1), 25–50.

Li, Z. (2002a). Pop-Out Theory: Segmentation Without Classification by the Primary Visual Cortex. In V. Cantoni, M. Marinaro, & A. Petrosino (Eds.), *Visual Attention Mechanisms* (pp. 69–78). Boston, MA: Springer US. https://doi.org/10.1007/978-1-4615-0111-4_7

Li, Z. (2002b). Saliency And Figure-Ground Effects. In V. Cantoni, M. Marinaro, & A. Petrosino (Eds.), *Visual Attention Mechanisms* (pp. 115–124). Boston, MA: Springer US. https://doi.org/10.1007/978-1-4615-0111-4_11

Li, Z. (2003). V1 mechanisms and some figure-ground and border effects. *Journal of Physiology Paris*, *97*(4), 503–515. https://doi.org/10.1016/j.jphysparis.2004.01.008

Liversedge, S. P., & Findlay, J. M. (2000). Saccadic eye movements and cognition. *Trends in Cognitive Sciences*, *4*(1), 6–14. https://doi.org/10.1016/S1364-6613(99)01418-7

Mackworth, N. H., & Thomas, E. L. (1962). Head-Mounted Eye-Marker Camera. *Journal of the Optical Society of America*, *52*(6), 713. https://doi.org/10.1364/JOSA.52.000713

Macmillan, N. A., & Creelman, C. D. (2005). Detection theory: A user's guide (2nd ed.). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Malik, J., & Perona, P. (1990). Preattentive texture discrimination with early vision mechanisms. *Journal of the Optical Society of America A*, *7*(5), 923–932. https://doi.org/10.1364/josaa.7.000923

Marino, R. A., & Munoz, D. P. (2009). The effects of bottom-up target luminance and top-down spatial target predictability on saccadic reaction times. *Experimental Brain Research*, *197*(4), 321–335. https://doi.org/10.1007/s00221-009-1919-x

McKendrick, A. M., Chan, Y. M., & Nguyen, B. N. (2018). Spatial vision in older adults: perceptual changes and neural bases. *Ophthalmic and Physiological Optics*, *38*(4), 363–375. https://doi.org/10.1111/opo.12565

Melcher, D., & Kowler, E. (1999). Shapes, surfaces and saccades. *Vision Research*, *39*(7), 2929–2946. https://doi.org/10.1016/S0042-6989(99)00029-2

Motoyoshi, I. (1999). Texture filling-in and texture segregation revealed by transient masking. *Vision Research*, *39*(7), 1285–1291. https://doi.org/10.1016/S0042-6989(98)00248-X

Motoyoshi, I., & Kingdom, F. A. A. (2010). The role of co-circularity of local elements in texture perception. *Journal of Vision*, *103*(1), 1–8. https://doi.org/10.1167/10.1.3

Norman, L. J., Heywood, C. A., & Kentridge, R. W. (2011). Contrasting the processes of texture segmentation and discrimination with static and phase-reversing stimuli. *Vision Research*, *51*, 2039–2047. https://doi.org/10.1016/j.visres.2011.07.021

Norton, T. T., Corliss, D. A., & Bailey, J. E. (2002). *The psychophysical measurement of visual function*. Boston, Mass: Butterworth-Heinemann.

Nothdurft, H. C. (1985a). Sensitivity for structure gradient in texture discrimination tasks. *Vision Research*, *25*(12), 1957–1968. https://doi.org/10.1016/0042-6989(85)90020-3

Nothdurft, H. C. (1985b). Orientation sensitivity and texture segmentation in patterns with different line orientation. *Vision Research*, *25*(4), 551–560. https://doi.org/10.1016/0042-6989(85)90159-2

Nothdurft, H. C. (1991). Texture segmentation and pop-out from orientation contrast. *Vision Research*, *31*(6), 1073–1078. https://doi.org/10.1016/0042-6989(91)90211-M

Nothdurft, H. C. (2000). Salience from feature contrast: temporal properties of saliency mechanisms. *Vision Research*, *40*(18), 2421–2435.

Nothdurft, H. C., & Parlitz, D. (1993). Absence of express saccades to texture or motion defined targets. *Vision Research*, *33*(10), 1367–1383. https://doi.org/10.1016/0042-6989(93)90044-W

Owsley, C. (2011). Aging and vision. *Vision Research*, *51*(13), 1610–1622. https://doi.org/https://doi.org/10.1016/j.visres.2010.10.020

Owsley, C., Sekuler, R., & Siemsen, D. (1983). Contrast sensitivity throughout adulthood. *Vision Research*. https://doi.org/10.1016/0042-6989(83)90210-9

Owsley, C., Burton, K.B. (1991). Aging and spatial contrast sensitivity: Underlying mechanisms and implications for everyday life. In Bagnoli, P., Hodow, W. (Eds.) The Changing Visual System (pp. 119-136). Boston, MA: Springer.

Palmer, S. E. (1999). *Vision science*. Massachusetts: MIT Press.

Paradiso, M. A., & Nakayama, K. (1991). Brightness perception and filling-in. *Vision Research*, *31*(7), 1221–1236. https://doi.org/10.1016/0042-6989(91)90047-9

Parkhurst, D. J., & Niebur, E. (2003). Scene content selected by active vision. *Spatial Vision*, *16*(2), 125–154. https://doi.org/10.1163/15685680360511645

Parkhurst, D. J., & Niebur, E. (2004). Texture contrast attracts overt visual attention in natural scenes. *European Journal of Neuroscience*, *19*, 783–789. https://doi.org/10.1111/j.1460-9568.2003.03183.x

Parkhurst, D., Law, K., & Niebur, E. (2002). Modeling the role of salience in the allocation of overt visual attention. *Vision Research*, *42*(1), 107–123. https://doi.org/10.1016/S0042-6989(01)00250-4

Pierce, J. (2007). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods*, *162*(1–2), 8–13.

Ploner, C. J., Ostendorf, F., & Dick, S. (2004). Target Size Modulates Saccadic Eye Movements in Humans. *Behavioral Neuroscience*. https://doi.org/10.1037/0735-7044.118.1.237

Pointer, J. S., & Hess, R. F. (1989). The contrast sensitivity gradient across the human visual field: With emphasis on the low spatial frequency range. *Vision Research*, *29*(9), 1133–1151. https://doi.org/10.1016/0042-6989(89)90061-8

Poort, J., Raudies, F., Wannig, A., Lamme, V. A. F., Neumann, H., & Roelfsema, P. R. (2012). The role of attention in figure-ground segregation in areas V1 and V4 of the visual cortex. *Neuron*, *75*, 143–156. https://doi.org/10.1016/j.neuron.2012.04.032

Posner, M. I., & Cohen, Y. (1984). Components of visual orienting. *Attention and Performance: Control of Language Processes*, 531–556. https://doi.org/10.1162/jocn.1991.3.4.335

Purves, D., Augustine, G., Fitzpatrick, D., Hall, W., LaMantia, A., & White, L. (Eds.). (2012). *Neuroscience* (5th ed.). Sunderland, MA.: Sinauer Associates, Inc.

Rayner, K. (1998). Eye movements in Reading and Information Processing: 20 Years of Research. *Psychological Bulletin*, *124*(3), 372–422. https://doi.org/10.1037/0033-2909.124.3.372

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology*. https://doi.org/10.1080/17470210902816461

Rayner, K., Rotello, C. M., Stewart, A. J., Keir, J., & Duffy, S. A. (2001). Integrating text and pictorial information: Eye movements when looking at print advertisements. *Journal of Experimental Psychology: Applied*, *7*(3), 219–226. https://doi.org/10.1037/1076-898X.7.3.219

Robol, V., Grassi, M., & Casco, C. (2013). Contextual influences in texture-segmentation: Distinct effects from elements along the edge and in the texture-region. *Vision Research*, *88*, 1–8. https://doi.org/10.1016/j.visres.2013.05.010

Rodieck, R. W. (1998). *The first steps in seeing*. Sunderland, MA,: Sinauer Associates.

Roelfsema, P. R., Lamme, V. A. F., Spekreijse, H., & Bosch, H. (2002). Figure – Ground Segregation in a Recurrent Network Architecture. *Journal of Cognitive Neuroscience*, *14*(4), 525–537. https://doi.org/10.1162/08989290260045756

Rossi, A. F., Desimone, R., & Ungerleider, L. G. (2001). Contextual Modulation in Primary Visual Cortex of Macaques. *The Journal of Neuroscience*, *21*(5), 1698–1709.

Schmolesky, M. T., Wang, Y., Pu, M., & Leventhal, A. G. (2000). Degradation of stimulus selectivity of visual cortical cells in senescent rhesus monkeys. *Nature Neuroscience*, *3*(4), 384–390. https://doi.org/10.1038/73957

Scholte, H. S., Jolij, J., Fahrenfort, J. J., & Lamme, V. A. F. (2008). Feedforward and Recurrent Processing in Scene Segmentation: Electroencephalography and Functional Magnetic Resonance Imaging. *Journal of Cognitive Neuroscience*, *20*(11), 2097–2109.

Schor, C. M. (2011). Neural Control of Eye Movements. In L. Levin, S. Nilsson, J. Ver Hoeve, S. Wu, P. Kaufman, & A. Alm (Eds.), *Adler's physiology of the eye* (11th ed., pp. 220–242). Edinburgh: Saunders Elsevier.

Skalka, H. W. (1980). Effect of age on Arden grating acuity. *British Journal of Ophthalmology*, *64*(1), 21–23. https://doi.org/10.1136/bjo.64.1.21

Snowden, R. J., & Kavanagh, E. (2006). Motion perception in the ageing visual system: Minimum motion, motion coherence, and speed discrimination thresholds. *Perception*. https://doi.org/10.1068/p5399

Stürzel, F., & Spillmann, L. (2001). Texture fading correlates with stimulus salience. *Vision Research*, *41*, 2969 – 2977. https://doi.org/10.1016/S0042-6989(01)00172-9

Su, Y. R., He, Z. J., & Ooi, T. L. (2011). Seeing grating-textured surface begins at the border. *Journal of Vision*, *1114*(1), 1–14. https://doi.org/10.1167/11.1.14

Sumner, P. (2011). Determinants of saccade latency. In S. Liversedge, I. Gilchrist, & S. Everling (Eds.), *The Oxford Handbook of Eye Movements*. New York: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199539789.013.0022

Supèr, H. (2006). Figure-ground activity in V1 and guidance of saccadic eye movements. *Journal of Physiology Paris*, *100*(1–3), 63–69. https://doi.org/10.1016/j.jphysparis.2006.09.002

Supèr, H., & Lamme, V. A. F. (2007). Strength of figure-ground activity in monkey primary visual cortex predicts saccadic reaction time in a delayed detection task. *Cerebral Cortex*, *17*(6), 1468–1475. https://doi.org/10.1093/cercor/bhl058

Supèr, H., Spekreijse, H., & Lamme, V. A. F. (2001). Two distinct modes of sensory processing observed in monkey primary visual cortex (V1). *Nature Neuroscience*, *4*(3), 304–310. https://doi.org/10.1038/85170

Supèr, H., Spekreijse, H., & Lamme, V. A. F. (2003). Figure-ground activity in primary visual cortex (V1) of the monkey matches the speed of behavioral response. *Neuroscience Letters*, *344*(2), 75–78. https://doi.org/10.1016/S0304-3940(03)00360-4

Supèr, H., van der Togt, C., Spekreijse, H., & Lamme, V. A. F. (2003). Internal state of monkey primary visual cortex (V1) predicts figure-ground perception. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *23*(8), 3407–3414. https://doi.org/23/8/3407 [pii]

Sutter, A., Sperling, G., & Chubb, C. (1995). Measuring the spatial frequency selectivity of second-order texture mechanisms. *Vision research*, *35*(7), 915-924.

Tadin, D., & Blake, R. (2005). Motion perception getting better with age? *Neuron*, *45*(3), 325–327. https://doi.org/10.1016/j.neuron.2005.01.017

Tang, Y., & Zhou, Y. (2009). Age-related decline of contrast sensitivity for second-order stimuli: Earlier onset, but slower progression, than for first-order stimuli. *Journal of Vision*, *9*(7), 1–15. https://doi.org/10.1167/9.7.18

Tatler, B. W. (2007). The central fixation bias in scene viewing: Selecting an optimal viewing position independently of motor biases and image feature distributions. *Journal of Vision*, *7*(14), 1–17. https://doi.org/10.1167/7.14.4

Thielscher, A., Kolle, M., Neumann, H., Spitzer, M., & Gron, G. (2008). Texture segmentation in human perception: A combined modeling and fMRI study. *Neuroscience*, *151*, 730–736. https://doi.org/10.1016/j.neuroscience.2007.11.040

Thielscher, A., & Neumann, H. (2003). Neural mechanisms of cortico–cortical interaction in texture boundary detection: a modeling approach. *Neuroscience*, *122*(4), 921–939. https://doi.org/10.1016/j.neuroscience.2003.08.050

Thielscher, A., & Neumann, H. (2005). Neural mechanisms of human texture processing: Texture boundary detection and visual search. *Spatial Vision*, *18*(2), 227–257. https://doi.org/10.1163/1568568053320594

Thielscher, A., & Neumann, H. (2007). A computational model to link psychophysics and cortical cell activation patterns in human texture processing. *Journal of Computational Neuroscience*, *22*, 255–282. https://doi.org/10.1007/s10827-006-0011-9

Tobii Technology. (n.d.-a). How do Tobii Eye Trackers work? Retrieved from https://www.tobiipro.com/learn-and-support/learn/eye-tracking-essentials/how-do-tobii-eye-trackers-work/

Tobii Technology. (n.d.-b). Dark and bright pupil tracking. Retrieved from https://www.tobiipro.com/learn-and-support/learn/eye-tracking-essentials/what-is-dark-and-bright-pupil-tracking/

Tobii Technology. (n.d.-c). What happens during the eye tracker calibration. Retrieved from https://www.tobiipro.com/learn-and-support/learn/eye-tracking-essentials/what-happens-during-the-eye-tracker-calibration/

Tobii Technology. (n.d.-d). Eye tracker accuracy and precision. Retrieved from https://www.tobiipro.com/learn-and-support/learn/eye-tracking-essentials/what-affects-the-accuracy-and-precision-of-an-eye-tracker/

Tobii Technology. (2010a). Timing Guide for Tobii Eye Trackers and Eye Tracking Software. Retrieved from https://www.tobiipro.com/siteassets/tobii-pro/learn-and-support/design/eye-tracker-timing-performance/tobii-eye-tracking-timing.pdf/?v=1.0

Tobii Technology. (2010b). Tobii T/X series Eye Trackers. Retrieved from https://www.tobiipro.com/siteassets/tobii-pro/product-descriptions/tobii-pro-tx-product-description.pdf/?v=1.0

Tobii Technology. (2011). Accuracy and precision test method for remote eye trackers. Retrieved from https://www.tobiipro.com/siteassets/tobii-pro/learn-and-support/use/what-affects-the-performance-of-an-eye-tracker/tobii-test-specifications-accuracy-and-precision-test-method.pdf/?v=2.1.1

Tobii Technology. (2012). Tobii Studio 3.2. Retrieved from https://www.tobiipro.com/siteassets/tobii-pro/product-descriptions/tobii-pro-studio-product-description.pdf/?v=3.2

Van Humbeeck, N., Schmitt, N., Hermens, F., Wagemans, J., & Ernst, U. A. (2013). The role of eye movements in a contour detection task. *Journal of Vision*, *13*(14), 1–19. https://doi.org/10.1167/13.14.5

Vancleef, K., Putzeys, T., Gheorghiu, E., Sassi, M., Machilsen, B., & Wagemans, J. (2013). Spatial arrangement in texture discrimination and texture segregation. *I-Perception*, *4*, 36–52. https://doi.org/10.1068/i0515

von Wartburg, R., Wurtz, P., Pflugshaupt, T., Nyffeler, T., Lüthi, M., & Müri, R. M. (2007). Size matters: Saccades during scene perception. *Perception*, *36*(3), 355–365. https://doi.org/10.1068/p5552

Wang, Z., Yu, S., Fu, Y., Tzvetanov, T., & Zhou, Y. (2018). Aging potentiates lateral but not local inhibition of orientation processing in primary visual cortex. *Frontiers in Aging Neuroscience*, *10*(14), 1–16. https://doi.org/10.3389/fnagi.2018.00014

Watson, A. B., & Pelli, D. G. (1983). Quest: A Bayesian adaptive psychometric method. *Perception & Psychophysics*, *33*(2), 113–120. https://doi.org/10.3758/BF03202828

Weigle, C., & Banks, D. (2008). Analysis of eye-tracking experiments performed on a Tobii T60. In *Proc. SPIE 6809, Visualization and Data Analysis 2008*. SPIE.

Wolfson, S. S., & Landy, M. S. (1995). Discrimination of orientation-defined texture edges. *Vision Research*, *35*(22), 2863–2877. https://doi.org/10.1016/0042-6989(94)00302-3

Wolfson, S. S., & Landy, M. S. (1998). Examining edge- and region-based texture analysis mechanisms. *Vision Research*, *38*(3), 439–446. https://doi.org/10.1016/S0042-6989(97)00153-3

Wood, J. M., & Bullimore, M. A. (1995). Changes in the lower displacement limit for motion with age. *Ophthalmic and Physiological Optics*, *15*(1), 31–36. https://doi.org/10.1016/0275-5408(95)92789-H

Yarbus, A. L. (1967). Eye Movements During Perception of Complex Objects. In *Eye Movements and Vision* (pp. 171–211). Boston, MA: Springer. https://doi.org/10.1007/978-1-4899-5379-7_8

Yokota, M., & Yokota, Y. (2007). Influence of small eye movement on perceptual filling-in time. In *2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (pp. 1550–1553). IEEE. https://doi.org/10.1109/IEMBS.2007.4352599

Yokota, M., & Yokota, Y. (2010). Eye movement inhibits the facilitation of perceptual filling-in. In *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (pp. 6629–6632). IEEE. https://doi.org/10.1109/IEMBS.2010.5627140

Ziebell, O., & Nothdurft, H. C. (1995). Facilitated detection of pop-out targets by localised cueing. *Perception*, *24*, 46–46.

Appendix A

Region of First Fixation Analysis for Eye Tracking Pilot Study (Chapter 8)

The region of first fixation (RFF) data is the proportion of first fixations (in percentages) to each AOI. A three-way (contrast × profile × region) ANOVA (see Figure A.1) was carried out on the RFF data. A significant main effect of Region was found, $F(1.40,32.17)=12.601$, $p<0.001$. First fixations are more often made to the Centre ($M=36.9\%$) and Edge ($M=40.3\%$) region compared to the Ground region ($M=22.7\%$). Main effects of Contrast and Profile, as well as the interaction between them were not reported. This is because for main effect analysis, the results are averaged across the levels of the other variables (i.e. for main effect of Contrast, the results are averaged across the Profiles and Region), which would results in 100% for all levels. This also applies to the 2-way interaction between Contrast and Profile (which is averaged across the Region variable), which yield 100% for all levels.

A significant Contrast x Region interaction [$F(2,46)=7.486$, $p=0.002$] shows that for the high contrast stimuli, the Centre and Edge region are fixated first more often compared to the Ground region ($M_{Centre}=43.70\%$, $M_{Edge}=38.37\%$, $M_{Ground}=17.93\%$). However, for the low contrast stimuli, the Edge region is fixated first more often compared to the Centre and Ground region ($M_{Centre}=30.11\%$, $M_{Edge}=42.33\%$, $M_{Ground}=27.56\%$). This indicates that participants are slightly better at locating the figure at first fixation when the contrast is higher.

A significant interaction between Profile and Region [$F(4,92)=7.221$, $p<0.001$] shows that participants first fixated at the Centre and Edge region more often then the Ground region for the Block ($M_{Centre}=44.07\%$, $M_{Edge}=40.68\%$, $M_{Ground}=15.25\%$) and Cornsweet ($M_{Centre}=35.59\%$, $M_{Edge}=44.07\%$, $M_{Ground}=20.43\%$) profile. However, this pattern of results is

not found for the Blur profile, as the proportion of first fixations to each region was not significantly different (M_{Centre}=31.23%, M_{Edge}=36.29%, M_{Ground}=32.48%). This implies that participants are worse at detecting the Blur figure at first fixation.

The interaction between Contrast, Profile and Region was significant, $F(4,92)=3.479$, $p=0.011$, with results indicating that participants' region of first fixation was influenced by the contrast of the profile for only the Block and Cornsweet profile, not the Blur profile. For the Block and Cornsweet profile, lower contrast stimuli have a decreased proportion of first fixations to the Centre region compared to the high contrast stimuli. However, the proportion of first fixations to the Edge region is not affected by contrast.

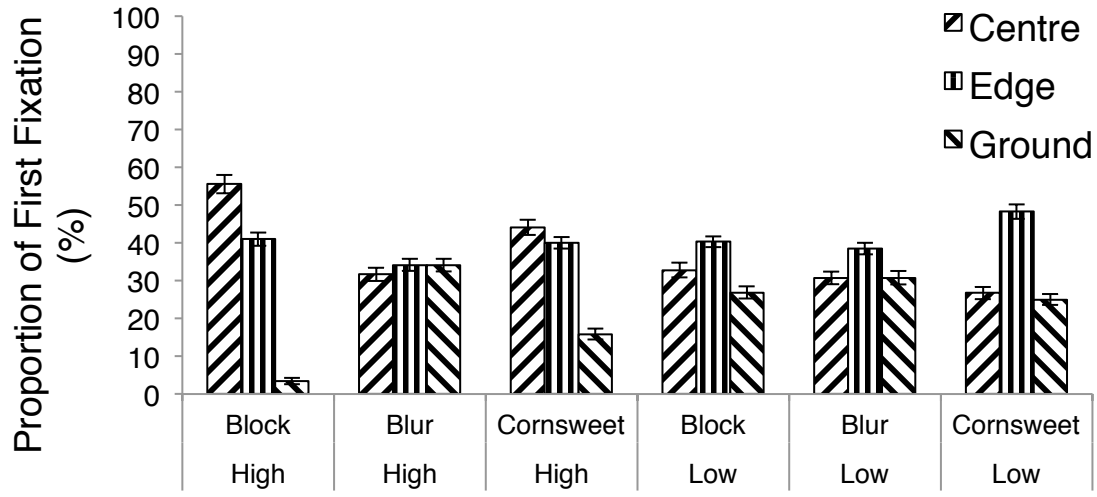
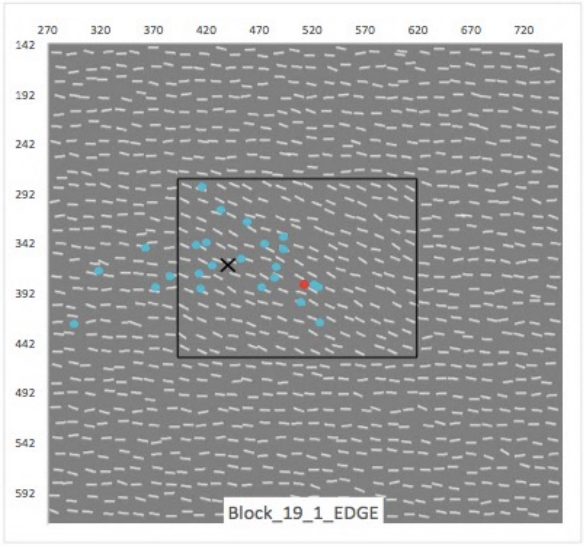


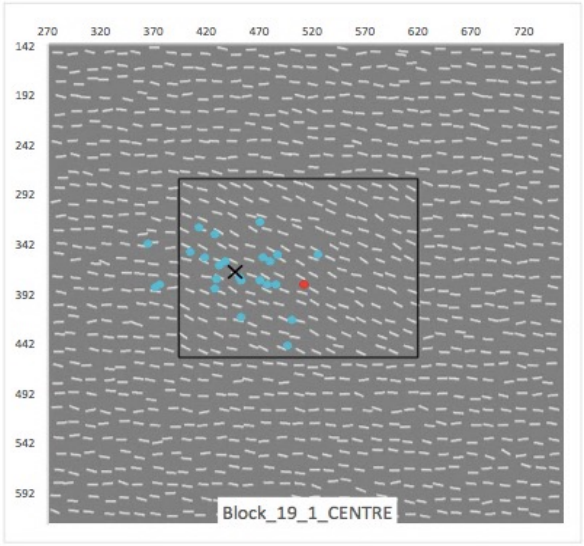
Figure A.1. Proportion of First Fixations to each AOI region for the various orientation contrasts and profiles. Error bars represent the standard error of the mean.

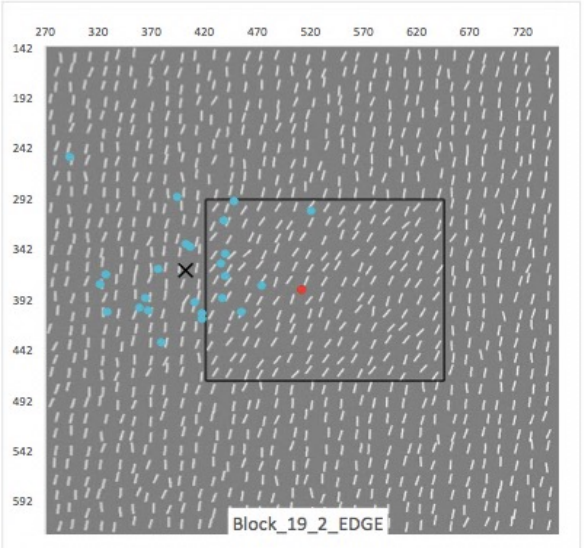
Appendix B

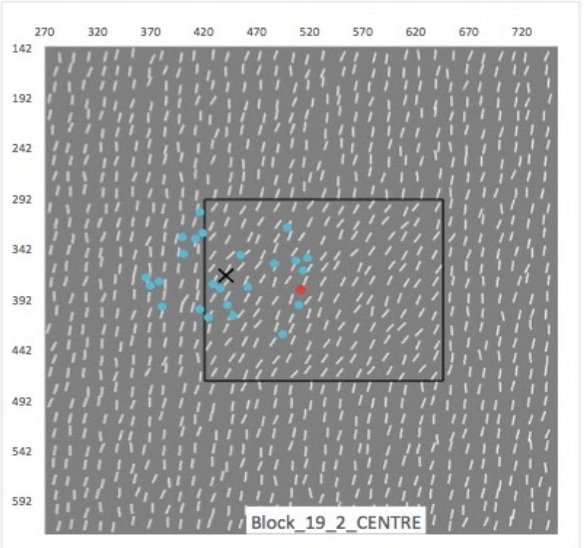
Position of First Fixation Plots for Experiment 7 (Chapter 8)

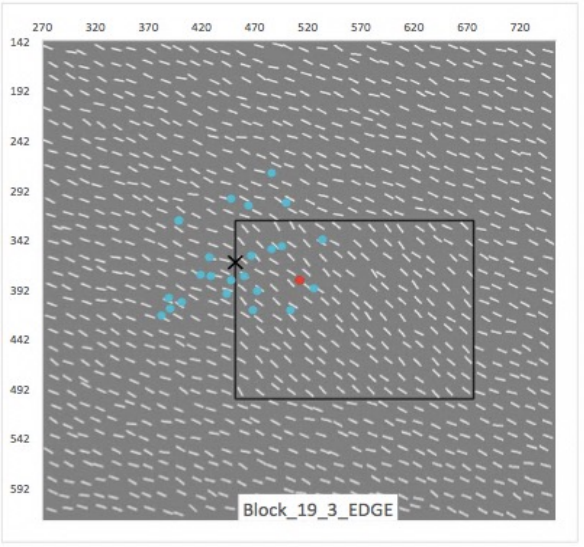
The plots here show the position of first fixations in terms of their xy-coordinates. The XY-axes represent the position of the texture grid in pixels. All 30 stimuli used in the task are shown here, with responses to the Edge task shown on the left column and responses to the Centre task shown on the right column. Each blue dot is one participants landing position, and the red dot is the centre of the 32×32 texture grid The black solid line on the stimuli represents the figure border, and the centroid (the mean landing position of first fixation for all participants) is represented by an ×.

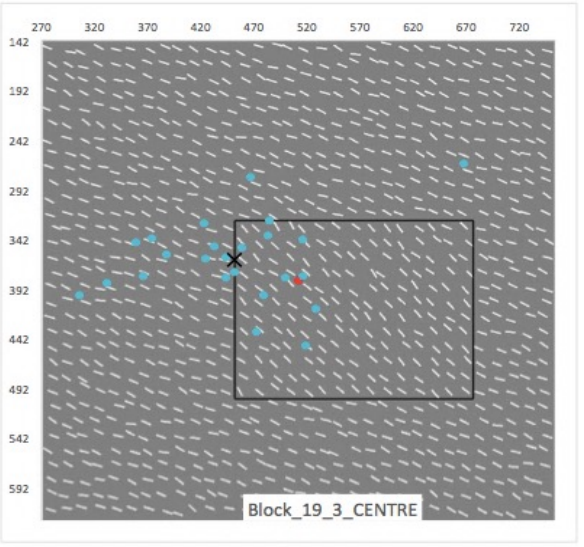

270 320 370 420 470 520 570 620 670 720
142 192 242 292 342 392 442 492 542 592
Block_19_1_EDGE

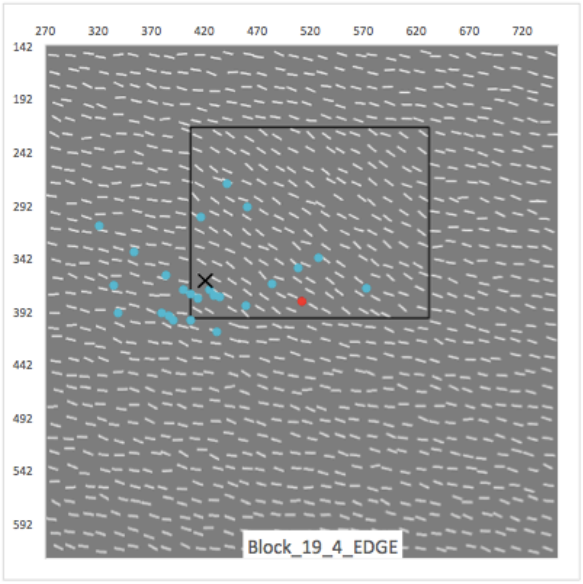

270 320 370 420 470 520 570 620 670 720
142 192 242 292 342 392 442 492 542 592
Block_19_1_CENTRE

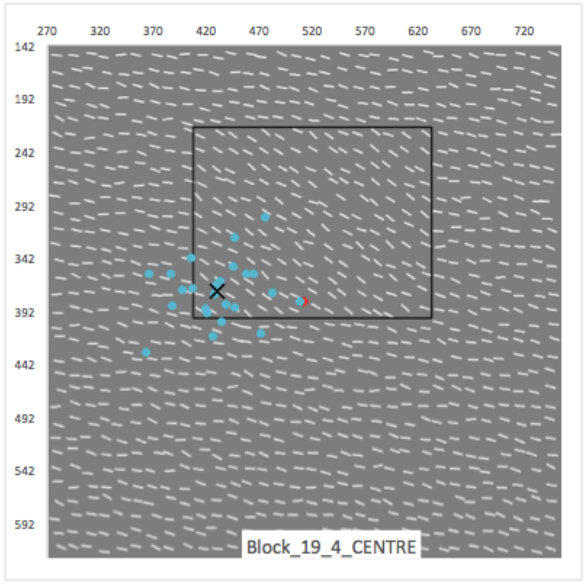

270 320 370 420 470 520 570 620 670 720
142 192 242 292 342 392 442 492 542 592
Block_19_2_EDGE


270 320 370 420 470 520 570 620 670 720
142 192 242 292 342 392 442 492 542 592
Block_19_2_CENTRE


270 320 370 420 470 520 570 620 670 720
142 192 242 292 342 392 442 492 542 592
Block_19_3_EDGE


270 320 370 420 470 520 570 620 670 720
142 192 242 292 342 392 442 492 542 592
Block_19_3_CENTRE


270 320 370 420 470 520 570 620 670 720
142 192 242 292 342 392 442 492 542 592
Block_19_4_EDGE


270 320 370 420 470 520 570 620 670 720
142 192 242 292 342 392 442 492 542 592
Block_19_4_CENTRE

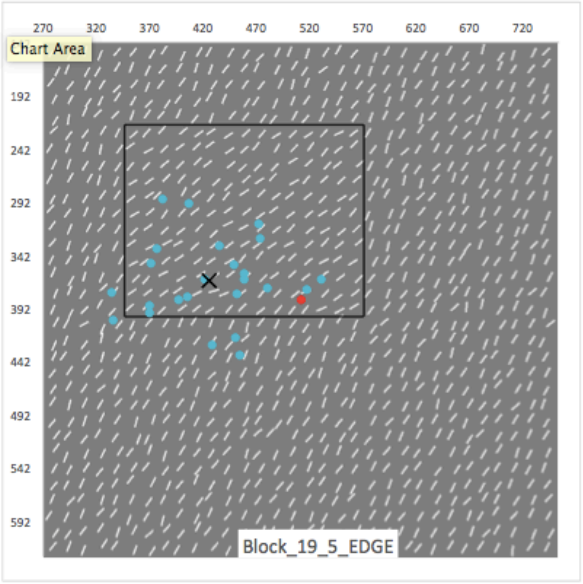
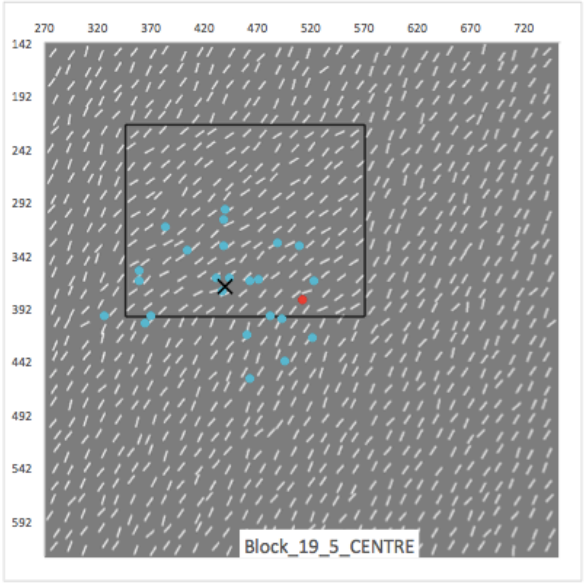


Chart Area
270 320 370 420 470 520 570 620 670 720
192 242 292 342 392 442 492 542 592
Block_19_5_EDGE


270 320 370 420 470 520 570 620 670 720
142 192 242 292 342 392 442 492 542 592
Block_19_5_CENTRE

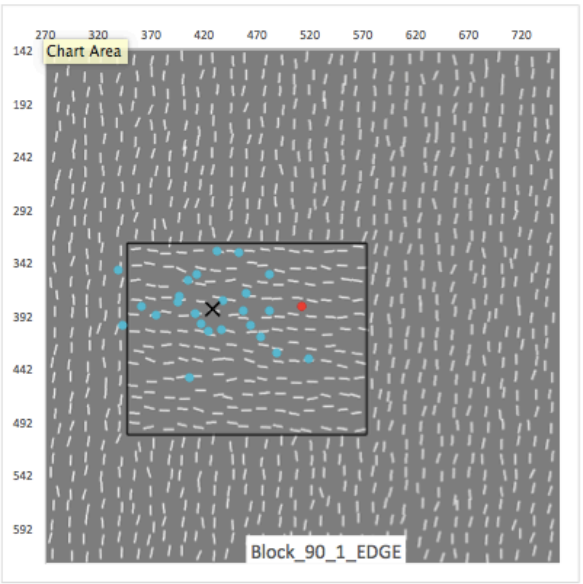
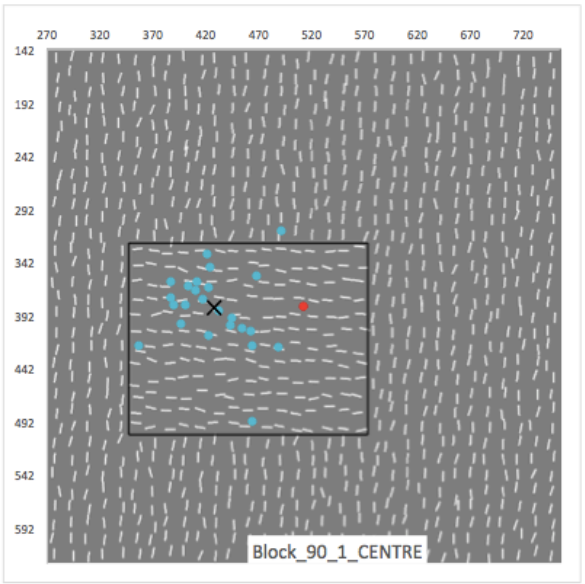
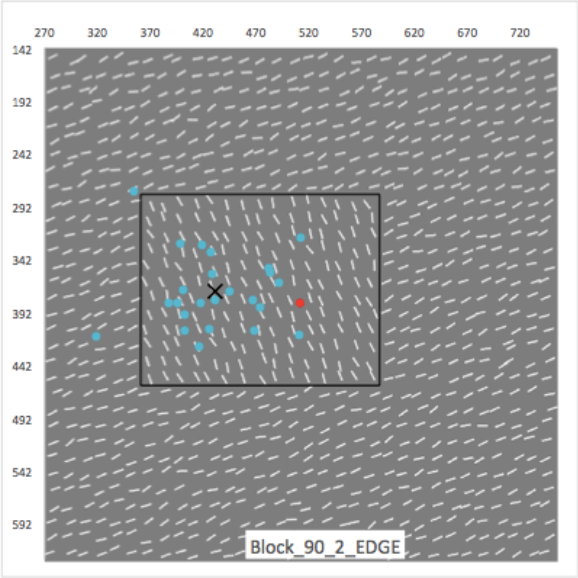
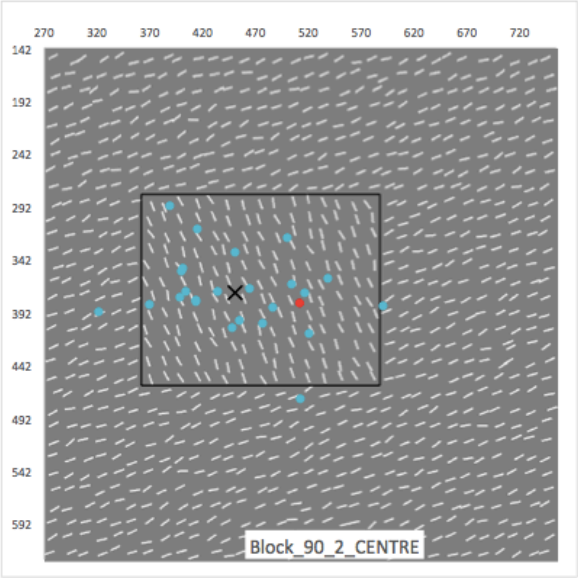
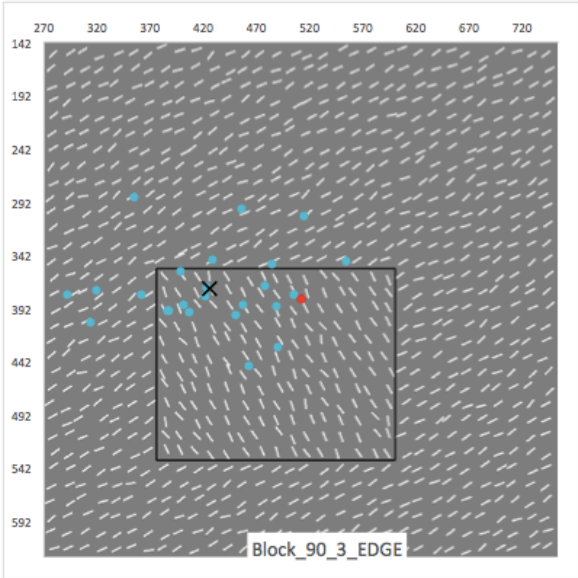


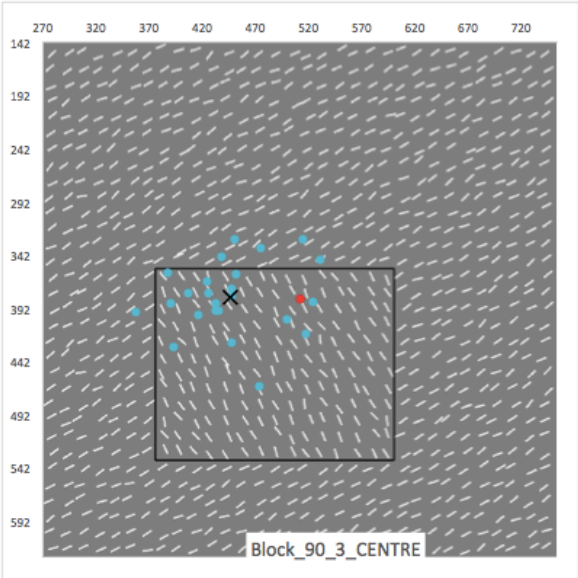
Chart Area
270 320 370 420 470 520 570 620 670 720
142 192 242 292 342 392 442 492 542 592
Block_90_1_EDGE

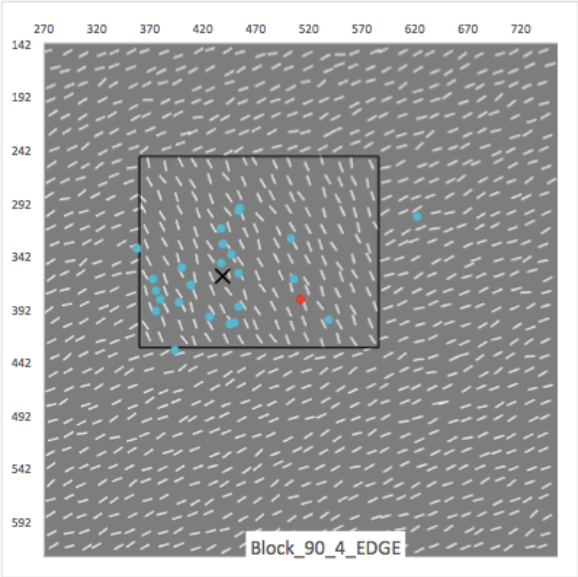

270 320 370 420 470 520 570 620 670 720
142 192 242 292 342 392 442 492 542 592
Block_90_1_CENTRE

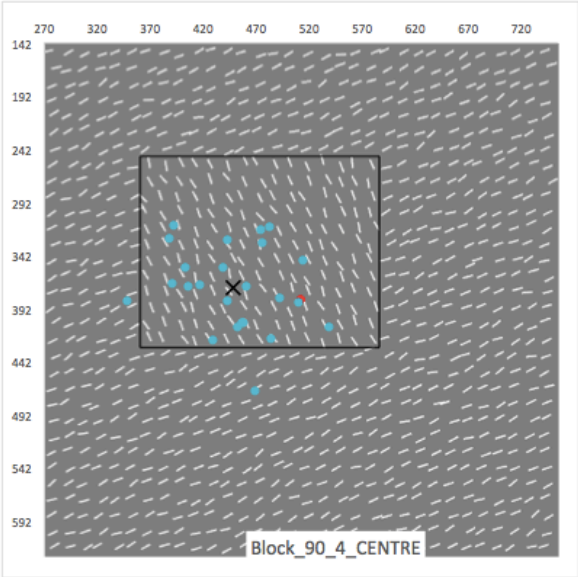

270 320 370 420 470 520 570 620 670 720
142 192 242 292 342 392 442 492 542 592
Block_90_2_EDGE

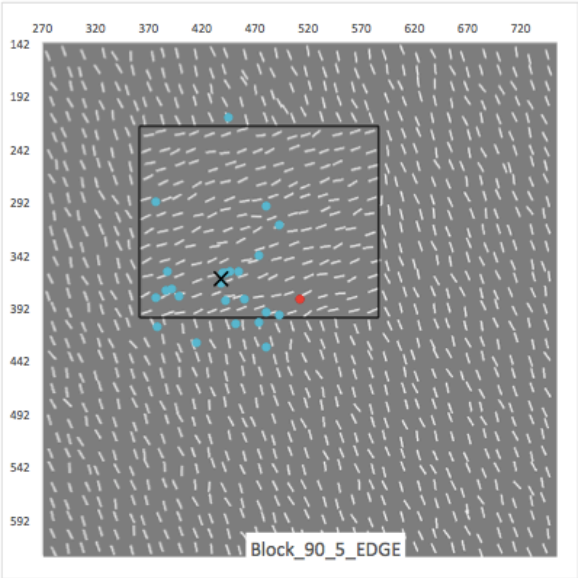

270 320 370 420 470 520 570 620 670 720
142 192 242 292 342 392 442 492 542 592
Block_90_2_CENTRE

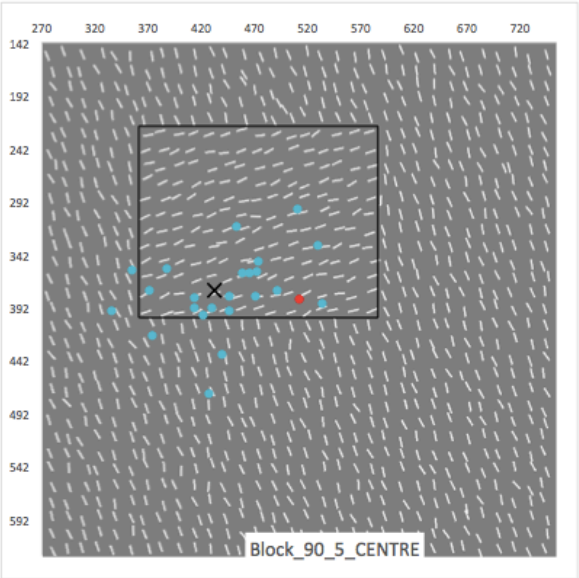

270 320 370 420 470 520 570 620 670 720
142 192 242 292 342 392 442 492 542 592
Block_90_3_EDGE

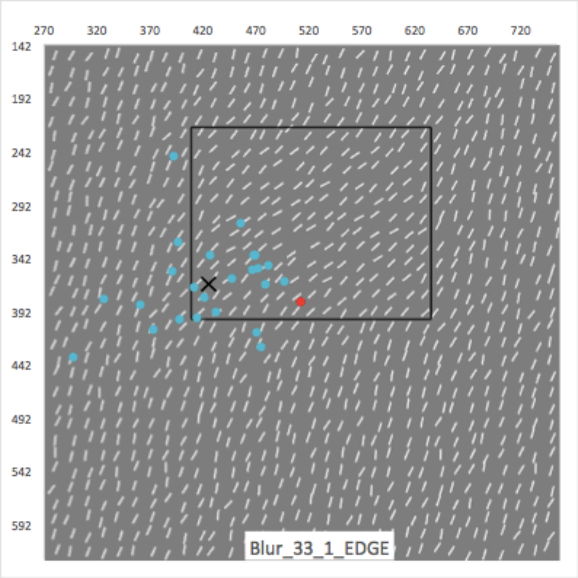

270 320 370 420 470 520 570 620 670 720
142 192 242 292 342 392 442 492 542 592
Block_90_3_CENTRE

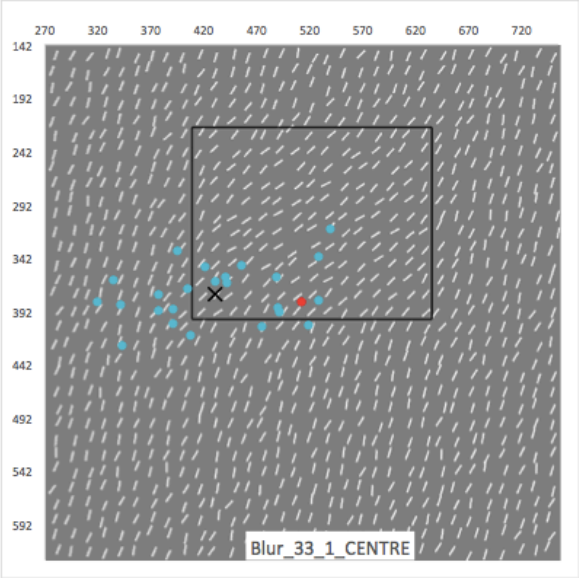

270 320 370 420 470 520 570 620 670 720
142 192 242 292 342 392 442 492 542 592
Block_90_4_EDGE

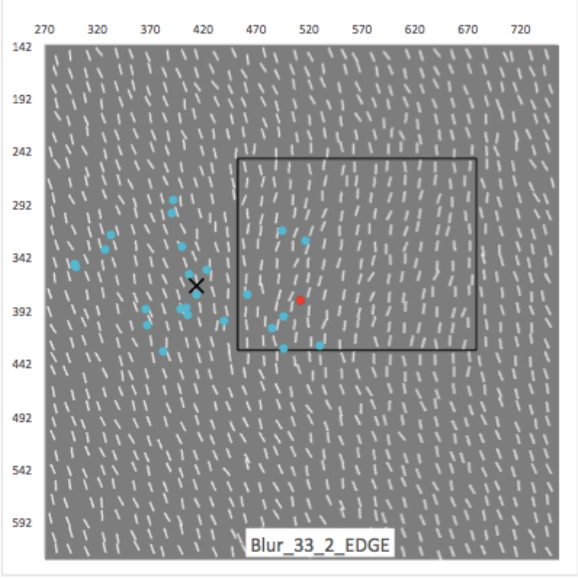

270 320 370 420 470 520 570 620 670 720
142 192 242 292 342 392 442 492 542 592
Block_90_4_CENTRE

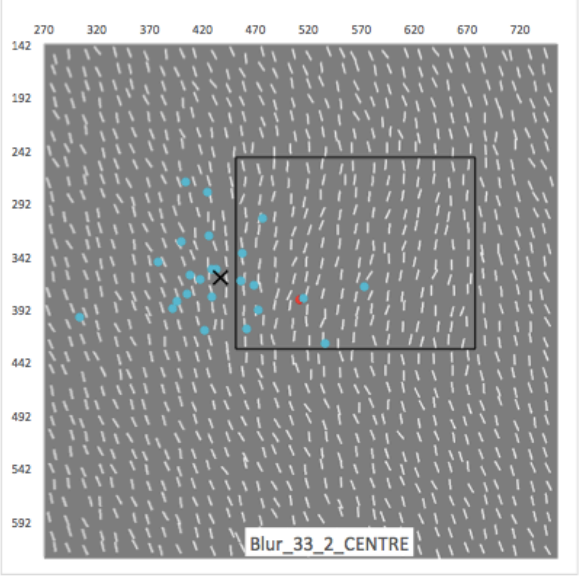

270 320 370 420 470 520 570 620 670 720
142 192 242 292 342 392 442 492 542 592
Block_90_5_EDGE

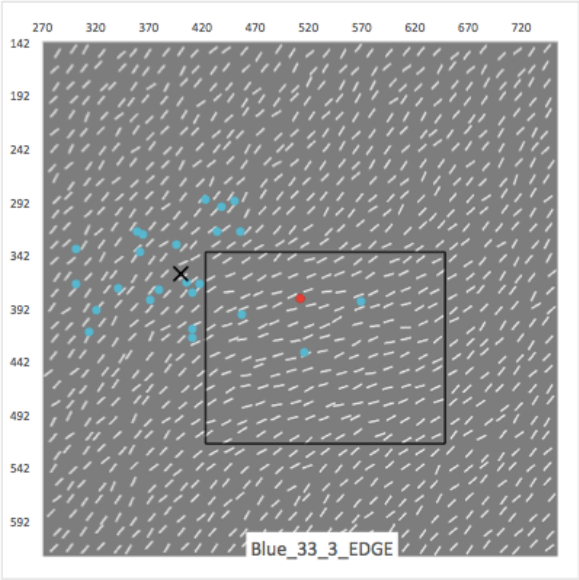

270 320 370 420 470 520 570 620 670 720
142 192 242 292 342 392 442 492 542 592
Block_90_5_CENTRE

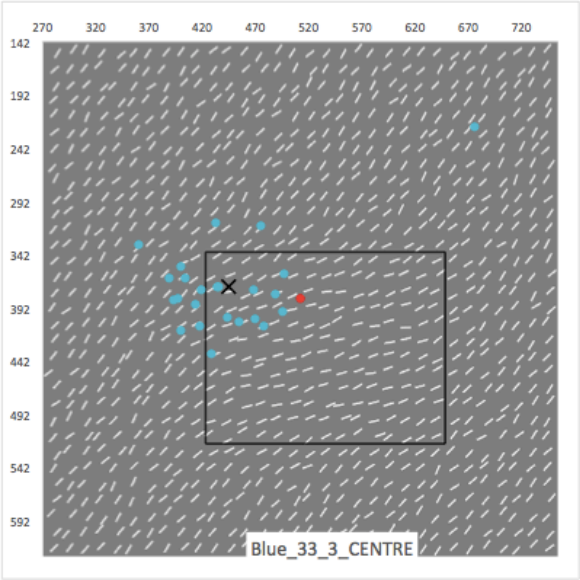

270 320 370 420 470 520 570 620 670 720
142 192 242 292 342 392 442 492 542 592
Blur_33_1_EDGE

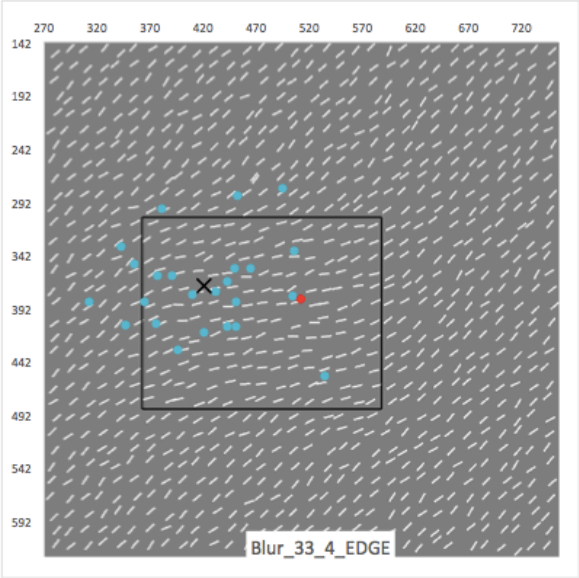

270 320 370 420 470 520 570 620 670 720
142 192 242 292 342 392 442 492 542 592
Blur_33_1_CENTRE

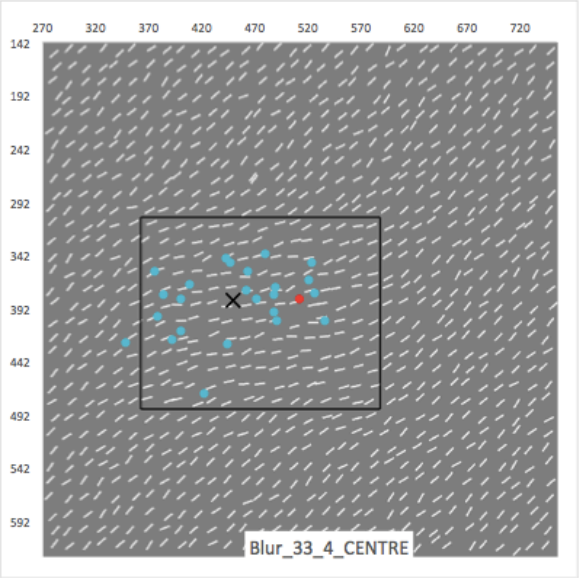

270 320 370 420 470 520 570 620 670 720
142 192 242 292 342 392 442 492 542 592
Blur_33_2_EDGE

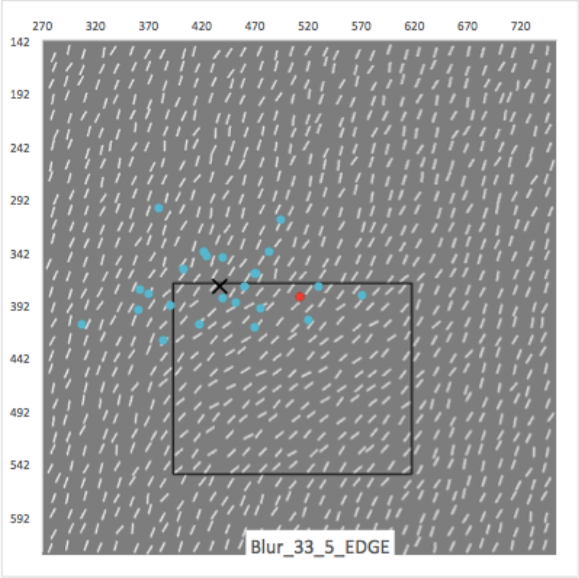

270 320 370 420 470 520 570 620 670 720
142 192 242 292 342 392 442 492 542 592
Blur_33_2_CENTRE

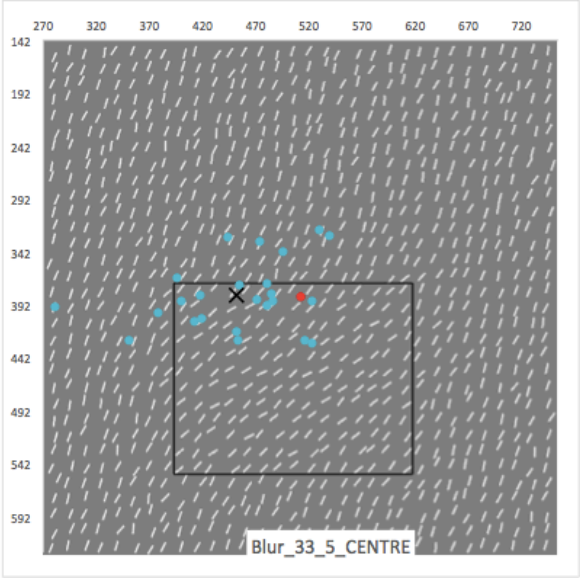

270 320 370 420 470 520 570 620 670 720
142 192 242 292 342 392 442 492 542 592
Blue_33_3_EDGE

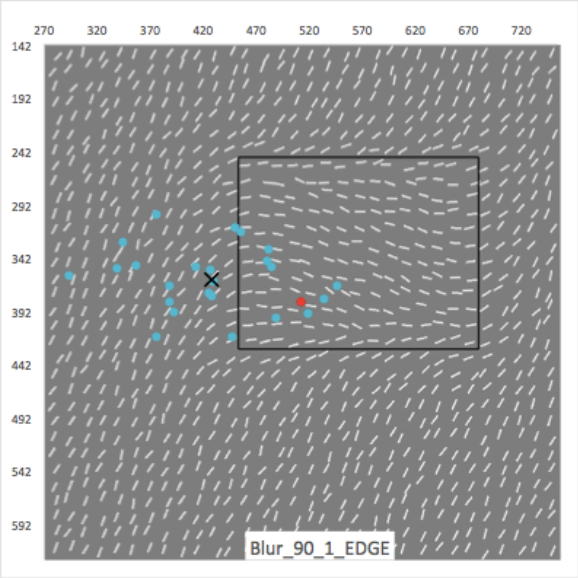

270 320 370 420 470 520 570 620 670 720
142 192 242 292 342 392 442 492 542 592
Blue_33_3_CENTRE

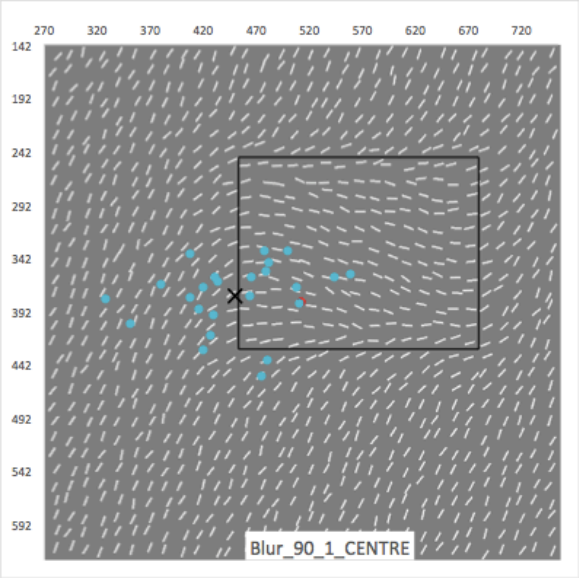

270 320 370 420 470 520 570 620 670 720
142 192 242 292 342 392 442 492 542 592
Blur_33_4_EDGE

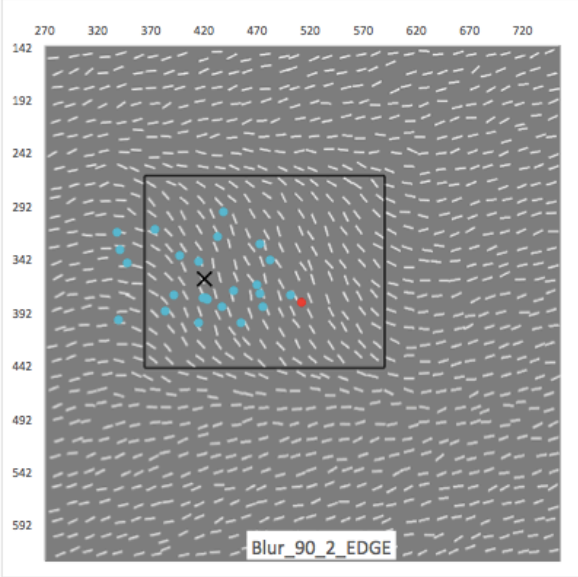

270 320 370 420 470 520 570 620 670 720
142 192 242 292 342 392 442 492 542 592
Blur_33_4_CENTRE

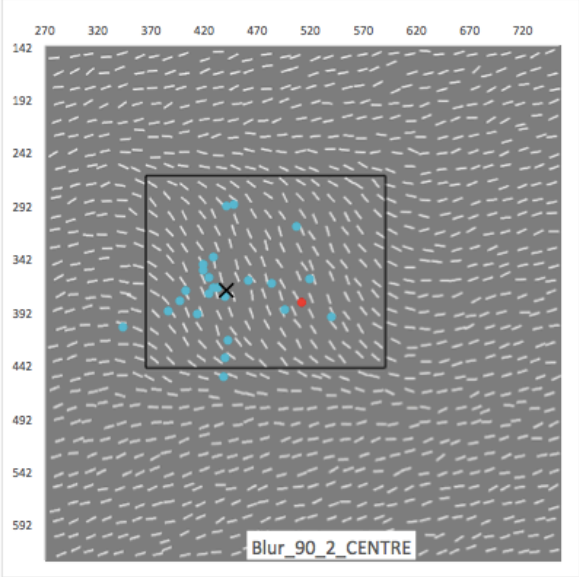

270 320 370 420 470 520 570 620 670 720
142 192 242 292 342 392 442 492 542 592
Blur_33_5_EDGE

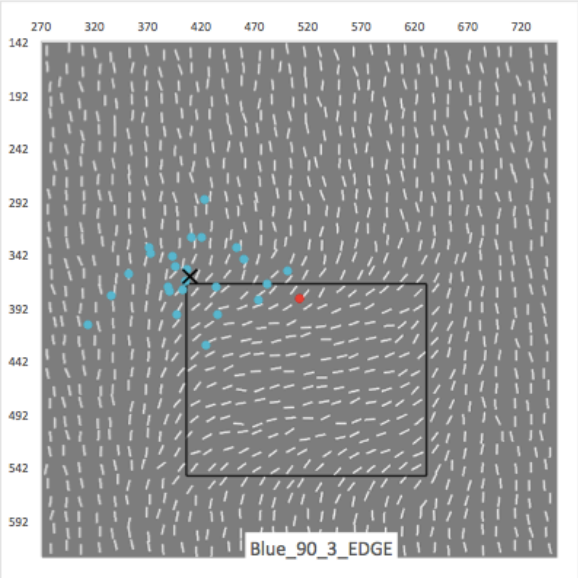

270 320 370 420 470 520 570 620 670 720
142 192 242 292 342 392 442 492 542 592
Blur_33_5_CENTRE

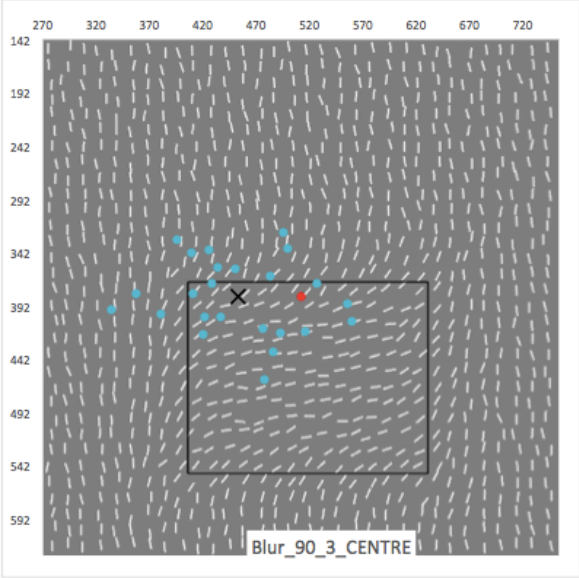

270 320 370 420 470 520 570 620 670 720
142 192 242 292 342 392 442 492 542 592
Blur_90_1_EDGE

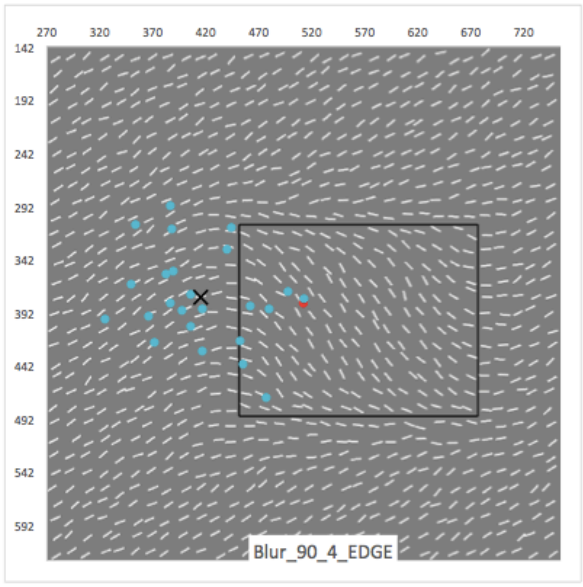

270 320 370 420 470 520 570 620 670 720
142 192 242 292 342 392 442 492 542 592
Blur_90_1_CENTRE

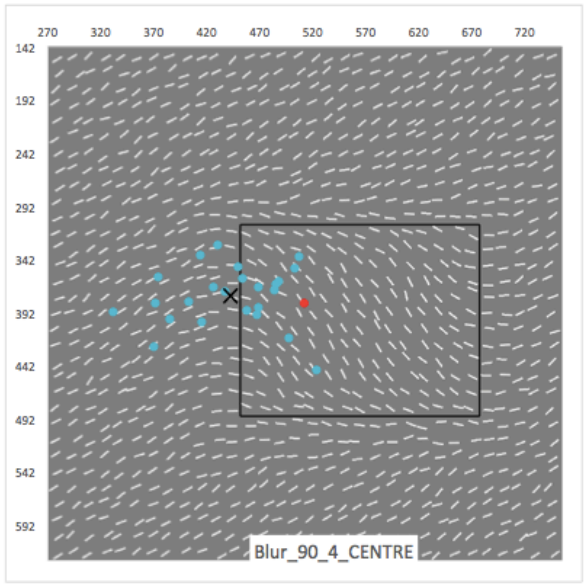

270 320 370 420 470 520 570 620 670 720
142 192 242 292 342 392 442 492 542 592
Blur_90_2_EDGE

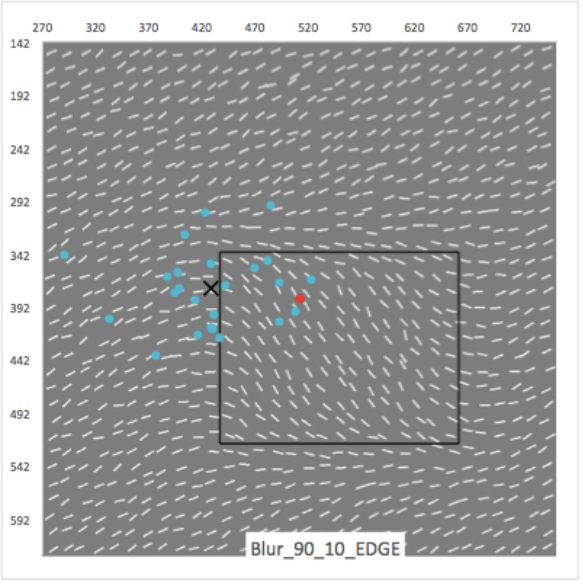

270 320 370 420 470 520 570 620 670 720
142 192 242 292 342 392 442 492 542 592
Blur_90_2_CENTRE

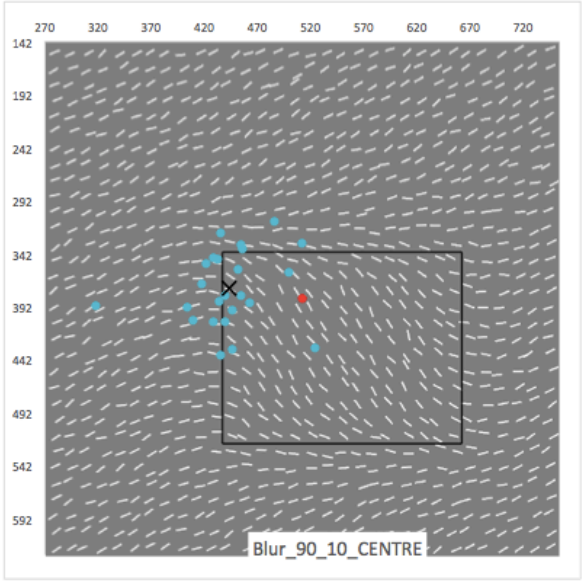

270 320 370 420 470 520 570 620 670 720
142 192 242 292 342 392 442 492 542 592
Blue_90_3_EDGE

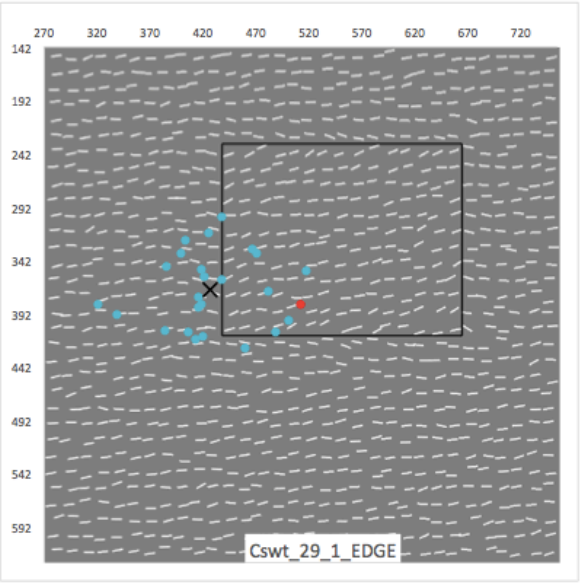

270 320 370 420 470 520 570 620 670 720
142 192 242 292 342 392 442 492 542 592
Blur_90_3_CENTRE

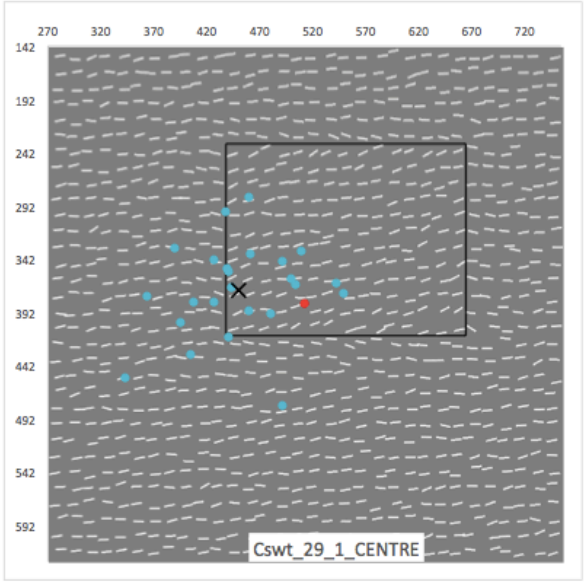

270 320 370 420 470 520 570 620 670 720
142 192 242 292 342 392 442 492 542 592
Blur_90_4_EDGE

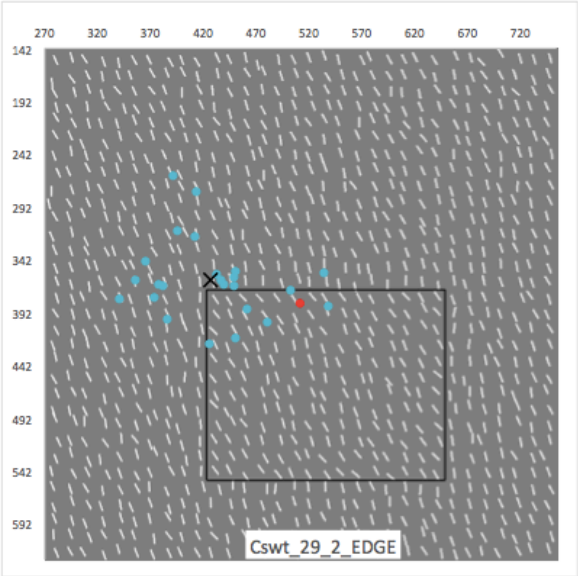

270 320 370 420 470 520 570 620 670 720
142 192 242 292 342 392 442 492 542 592
Blur_90_4_CENTRE

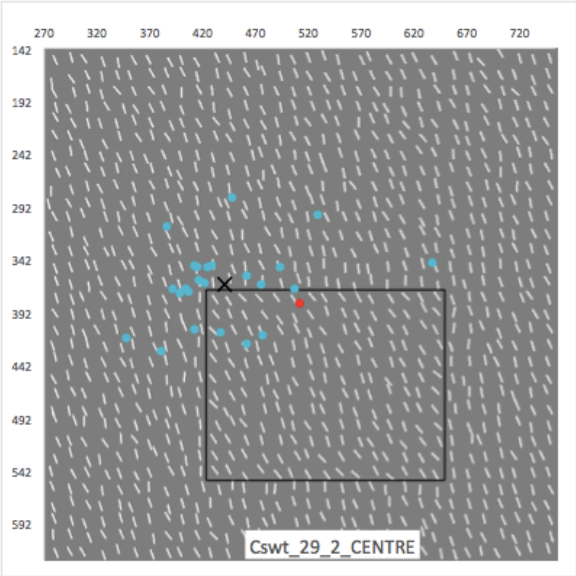

270 320 370 420 470 520 570 620 670 720
142 192 242 292 342 392 442 492 542 592
Blur_90_10_EDGE

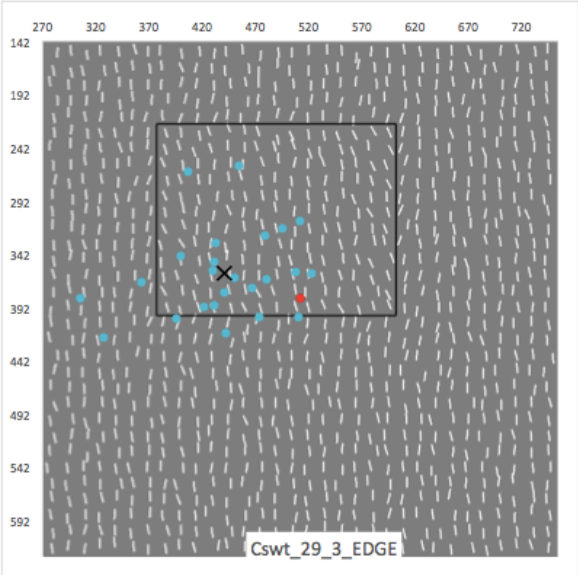

270 320 370 420 470 520 570 620 670 720
142 192 242 292 342 392 442 492 542 592
Blur_90_10_CENTRE

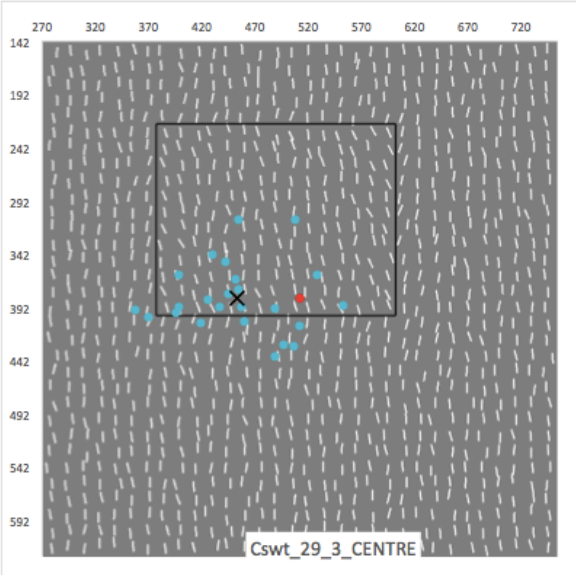

270 320 370 420 470 520 570 620 670 720
142 192 242 292 342 392 442 492 542 592
Cswt_29_1_EDGE

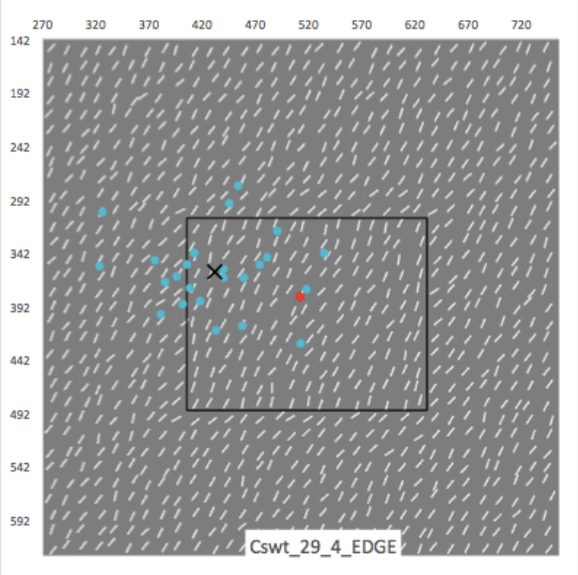

270 320 370 420 470 520 570 620 670 720
142 192 242 292 342 392 442 492 542 592
Cswt_29_1_CENTRE

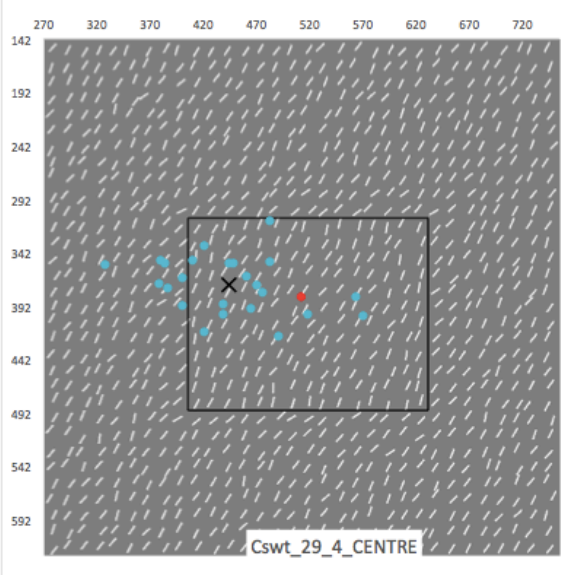

270 320 370 420 470 520 570 620 670 720
142 192 242 292 342 392 442 492 542 592
Cswt_29_2_EDGE

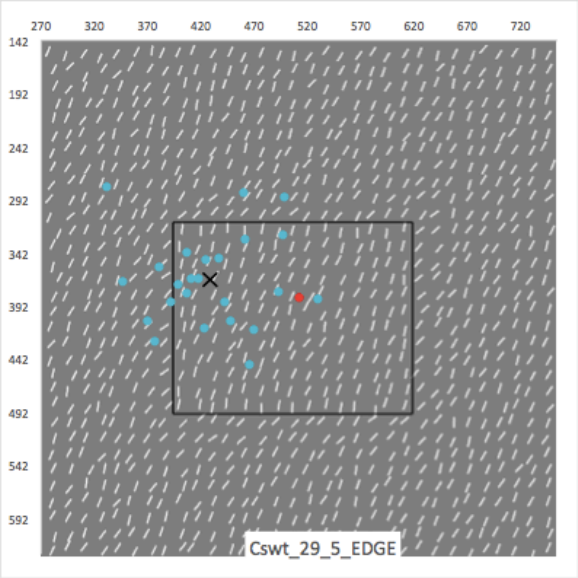

270 320 370 420 470 520 570 620 670 720
142 192 242 292 342 392 442 492 542 592
Cswt_29_2_CENTRE

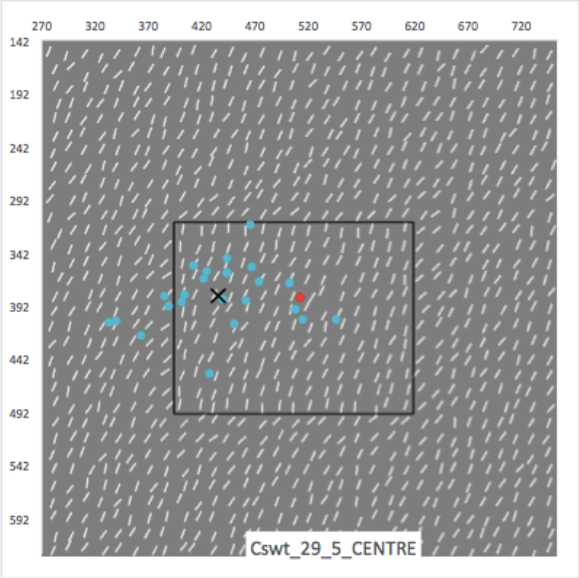

270 320 370 420 470 520 570 620 670 720
142 192 242 292 342 392 442 492 542 592
Cswt_29_3_EDGE

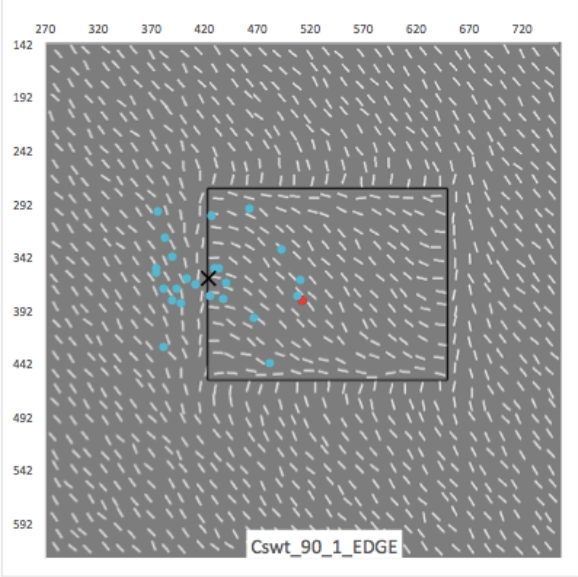

270 320 370 420 470 520 570 620 670 720
142 192 242 292 342 392 442 492 542 592
Cswt_29_3_CENTRE

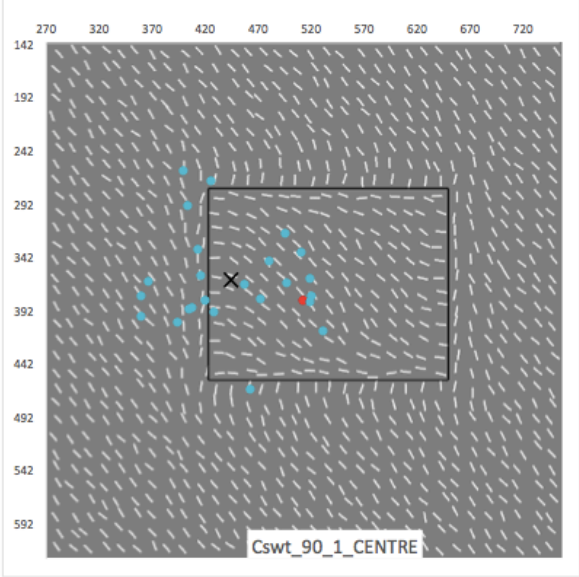

270 320 370 420 470 520 570 620 670 720
142 192 242 292 342 392 442 492 542 592
Cswt_29_4_EDGE

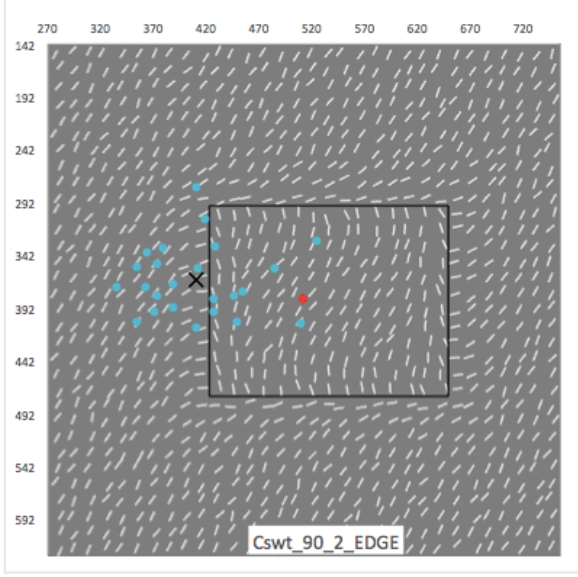

270 320 370 420 470 520 570 620 670 720
142 192 242 292 342 392 442 492 542 592
Cswt_29_4_CENTRE

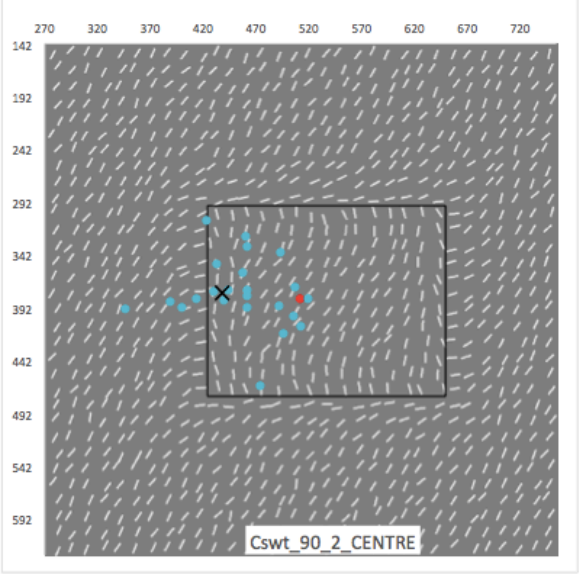

270 320 370 420 470 520 570 620 670 720
142 192 242 292 342 392 442 492 542 592
Cswt_29_5_EDGE


270 320 370 420 470 520 570 620 670 720
142 192 242 292 342 392 442 492 542 592
Cswt_29_5_CENTRE


270 320 370 420 470 520 570 620 670 720
142 192 242 292 342 392 442 492 542 592
Cswt_90_1_EDGE


270 320 370 420 470 520 570 620 670 720
142 192 242 292 342 392 442 492 542 592
Cswt_90_1_CENTRE


270 320 370 420 470 520 570 620 670 720
142 192 242 292 342 392 442 492 542 592
Cswt_90_2_EDGE


270 320 370 420 470 520 570 620 670 720
142 192 242 292 342 392 442 492 542 592
Cswt_90_2_CENTRE

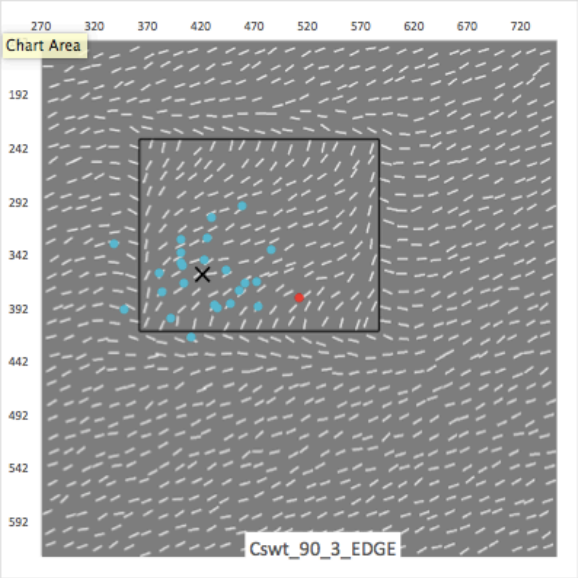
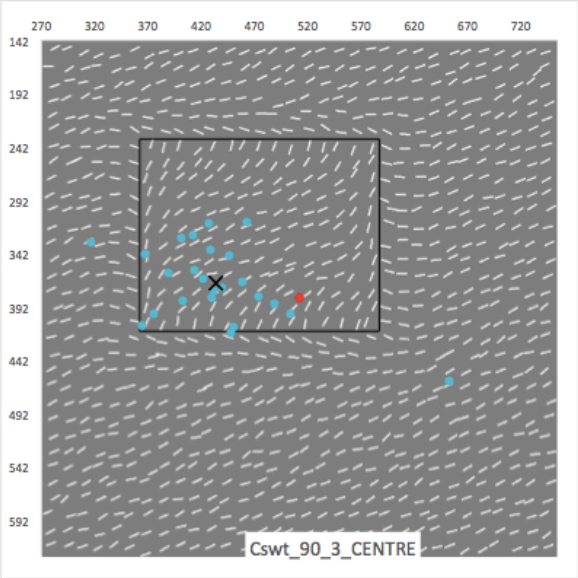
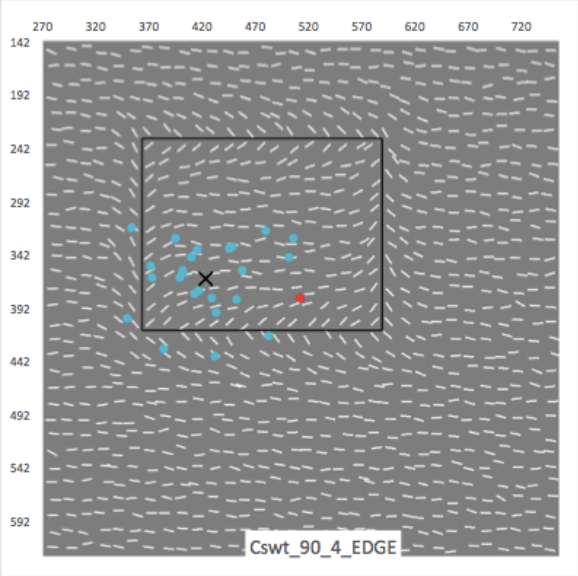
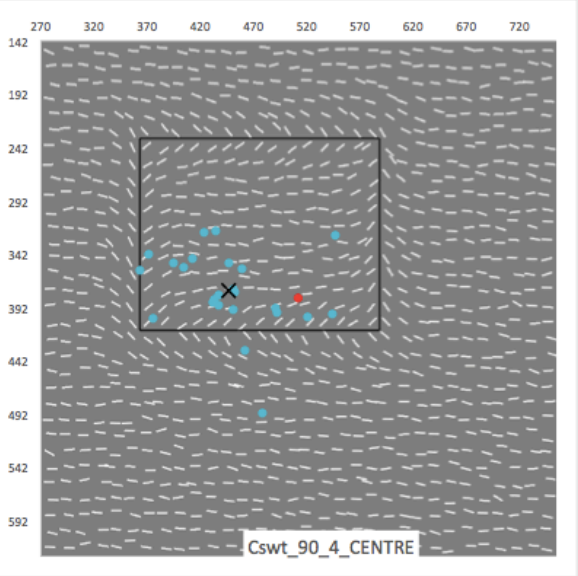
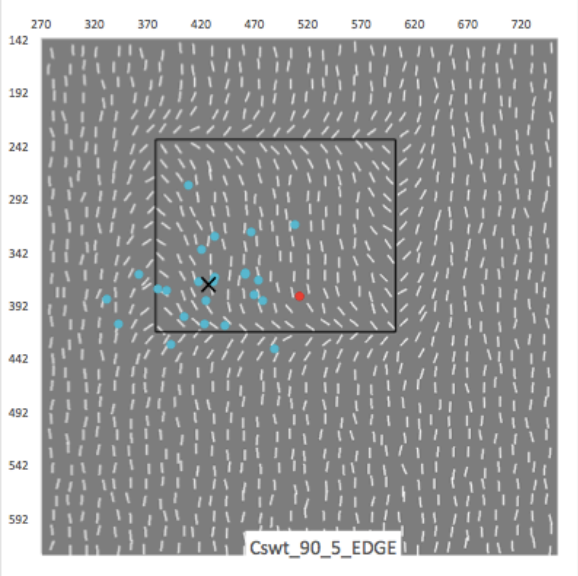


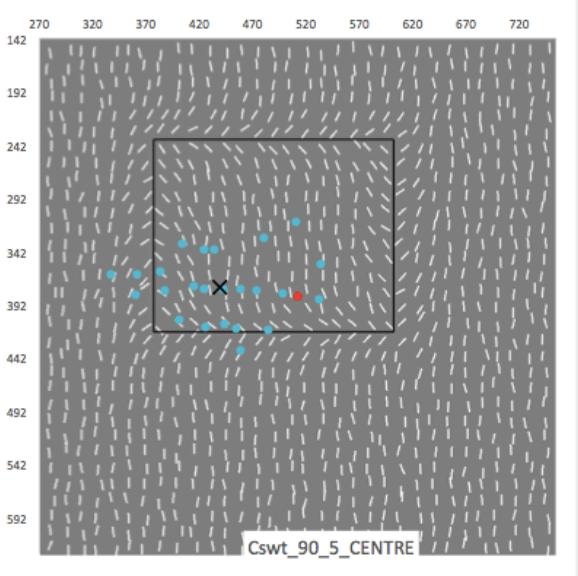
Chart Area
270 320 370 420 470 520 570 620 670 720
192 242 292 342 392 442 492 542 592
Cswt_90_3_EDGE


270 320 370 420 470 520 570 620 670 720
142 192 242 292 342 392 442 492 542 592
Cswt_90_3_CENTRE


270 320 370 420 470 520 570 620 670 720
142 192 242 292 342 392 442 492 542 592
Cswt_90_4_EDGE


270 320 370 420 470 520 570 620 670 720
142 192 242 292 342 392 442 492 542 592
Cswt_90_4_CENTRE


270 320 370 420 470 520 570 620 670 720
142 192 242 292 342 392 442 492 542 592
Cswt_90_5_EDGE


270 320 370 420 470 520 570 620 670 720
142 192 242 292 342 392 442 492 542 592
Cswt_90_5_CENTRE

Appendix C

Region of First Fixation Analysis for Experiment 7 (Chapter 8)

The region of first fixation (RFF) data is the proportion of first fixations (in percentages) to each AOI. Once again main effects of Contrast and Profile, as well as the interaction between them were not reported. This is because for main effect analysis, the results are averaged across the levels of the other variables (i.e. for main effect of Contrast, the results are averaged across the Profiles and Region), which would results in 100% for all levels. This also applies to the 2-way interaction between Contrast and Profile (which is averaged across the Region variable), which yield 100% for all levels.

Edge Task. A three-way (contrast × profile × region) ANOVA (see Figure C.1) was carried out on the RFF for the Edge Task. The ANOVA showed a significant main effect of Region [$F(2,42) = 9.918$, $p<0.001$]: where participants were more likely to fixate for the first time to the Centre (M=33.56%), and Edge (M=47.172%) region, compared to the Ground region (M=19.268%).

There was a significant interaction between Contrast and Region [$F(2,46)= 6.710$ p=0. 003]. For the high contrast, participants' RFF was higher for the Centre (M=40.25%) and Edge (M=40.66%) region compared to the Ground region (M=19.1 %), while for the low contrast, participants RFF was higher for the Edge (M=53.69%) region only, and lower for the Centre (M=26.87%) and Ground (M=19.44%) region.

The ANOVA also revealed a significant interaction between Profile and Region [$F(4,84)=$ 4.912, p=0. 001]. For the Blur profile, the Edge (M=45.76%) region is more often the RFF compared to the Centre (M=25.38%) region. While for the Block profile, the Centre

(*M*=40.49%) and Edge (*M*=48.47%) regions were more often the RFF compared to the Ground region (*M*=11.1%), and for the Cornsweet profile, the RFF was equally high for the Centre (*M*=34.81%) and Edge (*M*=47.31%) region, but the RFF for the Ground (*M*=17.88%) was lower compared to the Edge.

There was no significant interaction between Contrast, Profile and Region [$F(4,84)= 1.884$, $p=0.121$].

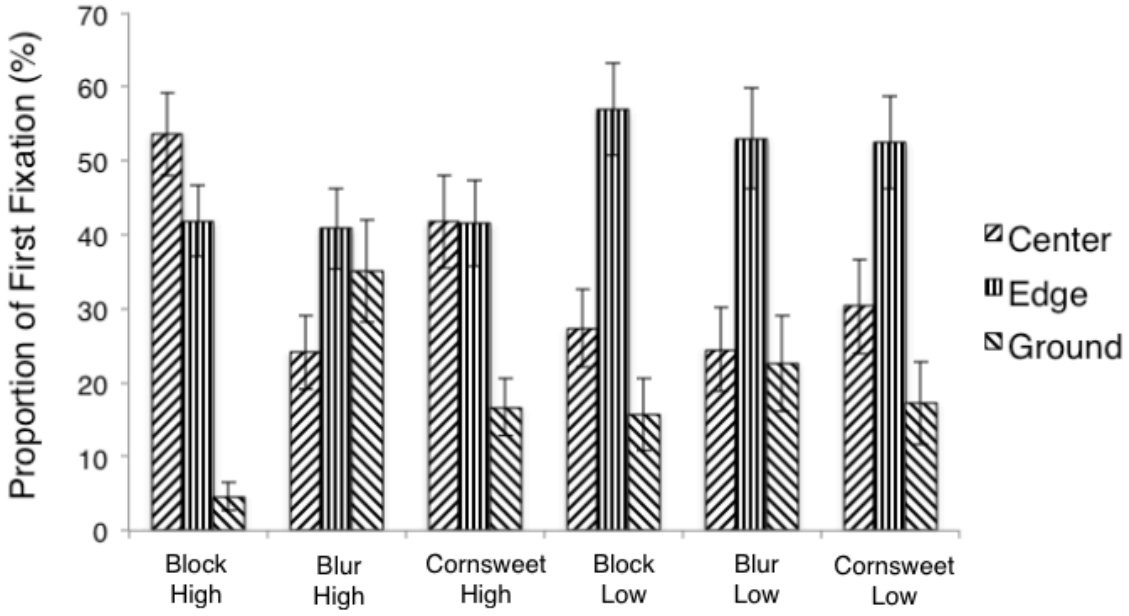


Figure C.1. Proportion of First Fixation to each region of the stimulus for the Edge task. Error bars represent the standard error of the mean.

Centre Task. A three-way (contrast × profile × region) ANOVA (see Figure C.2) was carried out on the RFF for the Centre Task. The ANOVA showed a significant main effect of Region [$F(2,42) = 45.616$, $p<0.001$]: where participants were more likely to fixate for the first time to the Edge (*M*=52.33%) region, followed by the Centre (*M*=39.84%) region, and least likely to the Ground region (*M*=9.84%).

The ANOVA also revealed a significant interaction between Profile and Region [$F(4,84)= 3.321$, $p=0.014$]. For the Block and Cornsweet profile, the Centre ($M_{Block}=44.13\%$; $M_{Cornsweet}= 40.56\%$) and Edge ($M_{Block}=48.76\%$; $M_{Cornsweet}= 52.26\%$) region were more often fixated on for the first time compared to the Ground region ($M_{Block}=7.08\%$; $M_{Cornsweet}= 7.18\%$). For the Blur profile on the other hand, the Edge ($M=55.94\%$) was fixated on first the most, followed by the Centre ($M=28.82\%$), and least to the Ground ($M=15.24\%$). There was no significant interaction between Contrast x Region [$F(2,46)= 2.702$ $p=0.078$], and Contrast x Profile x Region [$F(4,84)= 2.258$, $p=0.069$].

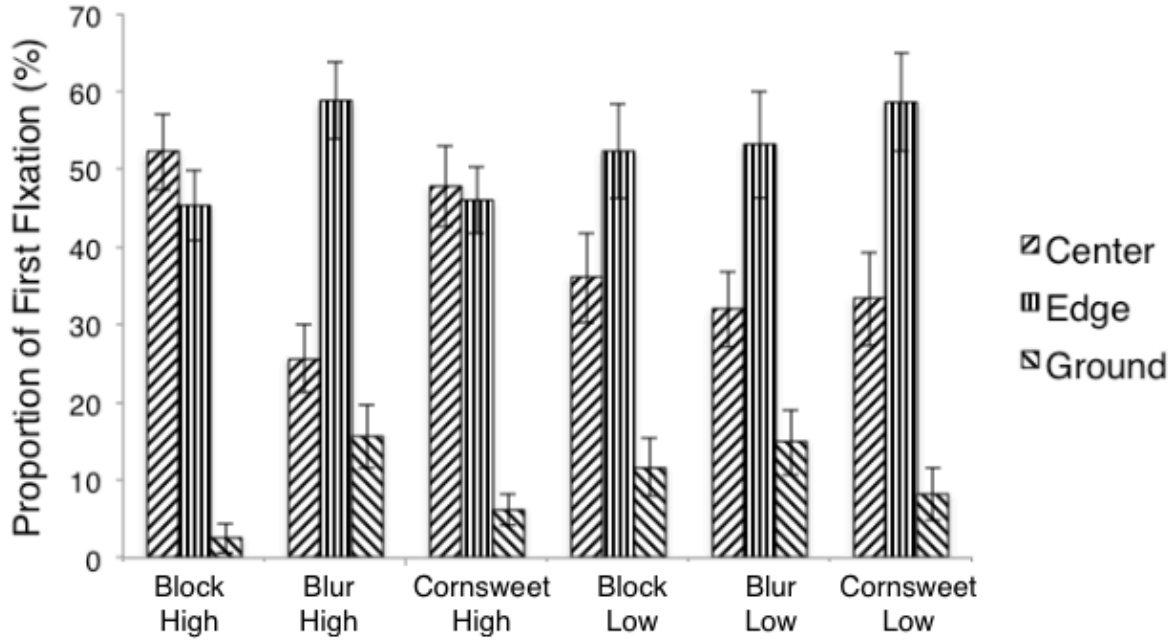


Figure C.2. Proportion of First Fixation to each region of the stimulus for the Centre task. Error bars represent the standard error of the mean.