Edge- and region-based processes of 2nd-order vision

LIAM NORMAN

How to cite:

NORMAN, LIAM (2013) Edge- and region-based processes of 2nd-order vision. Doctoral thesis, Durham University.

Edge- and region-based processes of 2nd-order vision

Liam J. Norman

The human visual system is sensitive to 2nd-order image properties (often called *texture* properties). Spatial gradients in certain 2nd-order properties are edge-based, in that contours are effortlessly perceived through a rapid segmentation process. Others, however, are region-based, in that they require regional integration in order to be discriminated. The five studies reported in this thesis explore these mechanisms of 2nd-order vision, referred to respectively as segmentation and discrimination. Study one compares the segmentation and discrimination of 2nd-order stimuli and uses flicker-defined-form to demonstrate that the former may be subserved by phase-insensitive mechanisms. In study two, through testing of a neuropsychological patient, it is shown that 2nd-order segmentation is achieved relatively early in the visual system and, contrary to some claims, does not require the region termed human "V4". Study three demonstrates, through selective adaptation aftereffects, that orientation variance (a 2nd-order regional property) is encoded by a dedicated mechanism tuned broadly to high and low variance and insensitive to low-level pattern information. Furthermore, the finding that the variance-specific aftereffect is limited to a retinotopic (not spatiotopic) reference frame, and that a neuropsychological patient with mid- to high-level visual cortical damage retains some sensitivity to variance, suggests that this regional property may be encoded at an earlier cortical site than previously assumed. Study four examines how cues from different 2nd-order channels are temporally integrated to allow cue-invariant segmentation. Results from testing a patient with bilateral lateral occipital damage and from selective visual field testing in normal observers suggest that this is achieved prior to the level of lateral occipital complex, but at least at the level of V2. The final study demonstrates that objects that are segmented rapidly by 2nd-order channels are processed at a sufficiently high cortical level as to allow object-based attention without those objects ever reaching awareness.

Edge- and region-based processes of 2nd-order vision

Norman, Liam John

Ph.D.

Department of Psychology

Durham University

2013

Contents

List of figures

List of tables

Declaration

I confirm that no part of the material offered has previously been submitted by me for a degree in this or any other University. If material has been generated through collaboration, my independent contribution has been clearly indicated. In all other cases material from the work of others has been acknowledged and quotations and paraphrases suitably indicated.

The entirety of chapter 2 is published in Vision Research:

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; doi: 10.1016/j.visres.2011.07.021

Section 11 of chapter 6 is published in Psychological Science:

Norman, L. J., Heywood, C. A., & Kentridge, R. W. (2013). Object-based attention without awareness. *Psychological Science, 24,* 836-843

Parts of chapter 6 were also presented at the 16th Annual Meeting for the Association for the Scientific Study of Consciousness:

16th Annual Meeting of the ASSC, Brighton, UK (2012). *Object-based attention is not sufficient for object awareness*

Copyright

Acknowledgments

I am extremely thankful to my supervisors Bob Kentridge and Charles Heywood, whose confidence in my abilities allowed all of the work reported in this thesis to happen. It has been a privilege to work with such dedicated researchers. In particular, I will forever be in debt to Bob (and Hannah Smithson) for introducing me to the world of computer programming for the purposes of vision science. Without these skills, the practical part of my PhD would have been a much more time-consuming (and less enjoyable) affair.

I also need to thank the participants in my experiments, namely those who took part in many of them over three years and often did so for no reward: Kam, Jamal, David, Manuela, Bobby and “Chemistry Dave”. Of course, I would also like to thank Michael and Debbie, who provided so much of their time and effort towards some of the parts in this thesis. I must also acknowledge Lee de-Wit, who provided a lot of fruitful discussion towards the experiments reported in chapter four, and who allowed me to collaborate on replicating an experiment of his which, unfortunately, only led to confusing (and somewhat paranormal) results. I am also very grateful to Bob Rafal, who provided me with a fantastic opportunity to attend the Visceral Mind summer school at Bangor in the later stages of my PhD.

As ever, my parents have been very supportive and continue to be so and I thank them dearly for that. Last of all, and certainly not least of all, I owe a great deal to Kam, who has shown incredible support throughout these 3 years and with whom I have shared many fond memories during my time as a PhD student. I hope that she will feel the same at the end of her PhD.

1. General introduction

Under normal conditions, visual perception rarely feels like anything less than a fully integrated experience. What is very clear from vision science, however, is that separate neural systems exist for the processing of many difference visual attributes. One such broad dissociation, often revealed subtly in carefully constructed experimental settings or strikingly in cases of selective neuropsychological disorders, concerns the respective processing of edge- and region-based visual information. With edge-based perception, the visual system is very sensitive to the locations at which two regions abut or overlap, provided that there is a coinciding discontinuity of some visual information. It is through this basic mechanism that the visual system achieves *segmentation* - an early representation of what might constitute figure and ground and which allows objects to be localised. In contrast, region-based analyses allow detailed descriptions of objects and surfaces to be formed and are thus essential for accurate object recognition. The term *discrimination* will often be used in this thesis to refer generally to these latter processes, primarily to distinguish them from those of segmentation.

The specific focus of this thesis is on the segmentation and discrimination of *2nd-order* visual attributes, which are often termed *texture* attributes. In contrast to 1st-order attributes, which can be represented by single-point analysis and are therefore limited to luminance and colour (Cavanagh & Mather, 1989), 2nd-order attributes cannot be detected in such a way by these linear mechanisms as they require analysis of multiple points in space (e.g. spatial frequency, orientation statistics, density, irregularity), but are nonetheless prevalent in natural scenes and important visual cues (Johnson & Baker, 2004; Schofield, 2000). Prior to describing in detail what is currently known about how the edge- and region-based properties of such cues are represented in human vision, it is necessary to first describe the early components of the visual system. The first few sections of this introductory chapter will therefore describe the early structural and perceptual components of the visual system, focussing on the dissociation between edge- and region-based processes (section **1.1**). There will be a particular emphasis on how contours are rapidly extracted and how regional information is subsequently “filled in” before describing how such characteristics apply to 2nd-order vision (section **1.2**). The specific direction and contents of this thesis are outlined in section **1.3**.

1.1. Early physiological and perceptual components of vision

1.1.1. From retina to cortex

Whatever the behavioural or perceptual output of a particular visual process is to be, all begins at the retina. It is the task of the photoreceptors (rods and cones), which lie in the outer layer of the retina, to absorb photons and convert that energy into electrical signals. The output of these cells is in the form of a graded signal, which can be passed onto horizontal or bipolar cells (Boycott & Wassle, 1991; Verweij, Dacey, Peterson & Buck, 1999). Horizontal cells offer lateral connections between photoreceptors whereas the bipolar cells synapse with the retinal *ganglion cells.* These cells are found in the more inner layers of the retina and collate the activity of varying numbers of photoreceptors, thus determining the spatial extent of their *receptive field.* The receptive field can be described as the area in visual space in which a stimulus elicits a measurable response in a cell's activity. Ganglion cells are not all the same, however, differing in their morphology to one another and reflecting different processing demands imposed upon the visual system.

The range of this functional and morphological diversity is reflected in the fact that an excess of 20 different ganglion cell types has been identified (Kolb, Linberg & Fisher, 1992). Only two types (*parasol* and *midget*), however, compose 80% of the entire population (Perry, Oehler & Cowey, 1984). Their names reflect the sizes of their cell bodies and dendritic trees, and thus the spatial extent of their receptive field. Parasol cells, for instance, are relatively large, receiving input from diffuse bipolar cells and extensive numbers of photoreceptors (Watanabe & Rodieck, 1989). Midget cells, conversely, receive input from very few midget bipolar cells and, within the region of the fovea, may innervate with as few as one, thus conserving spatial resolution as much as possible (Kolb & Dekorver, 1991; Kolb & Marshak, 2003). The ganglion cells converge and constitute the optic nerve, which resides at the back of the eye and exits through an opening in the retina. At the optic chiasm, projections from the nasal retinae in the optic nerve cross into contralateral hemispheres, whereas those from the temporal retinae stay on the same side. The optic tract then connects the chiasm to the dorsal lateral geniculate nucleus (dLGN) of the thalamus, wherein input from the two main ganglion cell types remains segregated. Parasol inputs converge on the 2 ventral magnocellular layers of the dLGN, whereas the midget inputs converge on the 4 dorsal parvocellular layers (Schiller & Malpelli, 1978; Shapley & Perry, 1986).

This segregation establishes the separate processing streams of the parvo- and magnocellular systems, which remain segregated at early stages of the visual system, each contributing to different functions. The cells of the magnocellular subdivision of visual processing (M cells) exhibit high temporal and low spatial sensitivity, as well as being adept at discriminating motion but being blind to colour identity. This is in stark contrast to the cells of the slower parvocellular system (P cells), which are more attuned to spatial detail and chromatic information (Schiller & Logothetis, 1990). In addition, another subtype of retinal ganglion cell, the bistratified type, forms the input to the konio cells of the dLGN, which are found between the magno- and parvocellular layers, as well as within them (Hendry & Reid, 2000). Cells within the dLGN retain the centre-surround antagonism of the retinal ganglion cells that precede them and from here many projections feed into the primary visual cortex in the occipital lobe, completing the retinogeniculostriate pathway.

1.1.2. Early cortical segregation of visual function

At the level of the primary visual cortex (V1, calcarine cortex, or Brodmann's area 17), concentrated at the calcarine sulcus, visual processing remains segregated to a large extent but quickly becomes increasingly complex. Six layers comprise V1, and the majority of the output from the dLGN congregates in the 4th layer. Within this layer, the sub-layer 4Cβ receives its input largely from parvocellular afferents and projects further to layer 3, whilst the sublayer 4Cα receives largely from magnocellular ones and projects to layer 4B (Blasdel & Lund, 1983; Hubel & Wiesel, 1972). Koniocellular afferents converge on layers 1-3 (Hendry & Yoshioka, 1994; Livingstone & Hubel, 1982).

Cytochrome oxidase (CO) labeling reveals an arrangement of two distinct cell types in the primary visual cortex: the "blobs" and the "interblobs". Both receive some input from layer 4Cβ (parvocellular) yet exhibit contrasting response properties to one another. Cells in the blobs show sensitivity to wavelength (colour) and brightness information but respond minimally to orientation information. Cells in the interblobs show the contrasting response profile. Furthermore, it has been suggested that the blobs are the neural bases of hue maps distributed throughout V1 (Lu & Roe, 2008; Xiao, Casti, Xiao & Kaplan, 2007). Additionally, although interblob cells are not selective for wavelength, some are sensitive to the presence of a contour defined purely by chromatic discontinuity (Dow, 2002; Livingstone & Hubel, 1988). From the blobs and interblobs, neural connections stem to the thin stripes and interstripes, respectively, of the second visual cortical area (V2). These anatomical distinctions are again revealed by CO labeling. Both pathways (blob -> thin stripe & interblob -> interstripes) project to the cortical area V4. Due to the relative response profiles and sensitivities of these pathways, some researchers have been led to claim that the blobs and interblobs in the primary visual cortex represent a distinction of “surface” and “contour” processing, respectively, in the early visual system (Livingstone & Hubel, 1988). The magnocellular subdivision of visual processing, contrarily, connects layers 4Cα and 4B in V1 and the CO-rich thick stripes of V2. From here, the magnocellular pathway connects largely, although not exclusively to, the dorsal stream of cortical visual processing.

1.1.3. Prioritising locations of contrast

As seen in the previous chapter, the hierarchical nature of the human visual system demands that we understand the mechanisms of higher-order visual processes (i.e. how objects are constructed) first through understanding those of lower levels that lay the foundations for all subsequent neural processes. One fundamental property of early visual processing, which was touched upon in the previous section, is the strong sensitivity to differential illumination in contrast to the weak sensitivity to diffuse illumination. This is reflected first in the receptive field structure of retinal ganglion cells, which have been the subject of much study since their initial description in the 1950s (Kuffler, 1953; Hartline and Ratliff, 1957; 1958). Two main types of receptive field structure, which usually adopt a circular shape, are common characteristics of the ganglion cell: the *on-centre* and *off-centre* receptive fields. In the former, maximum excitation of the ganglion cell is achieved when a bright circular stimulus is present in the central region of the receptive field, and is simultaneously surrounded by an annular dark patch. The converse is true for cells of an off-centre structure. Consequently, uniformly stimulating either type with a uniformly light or dark stimulus would not excite the cell’s activity. This selectivity renders diffuse regions of uniform luminance rather redundant at the earliest stages of the visual system, an aspect which lays the foundations for subsequent neural processing.

Within V1, for example, cells possess receptive fields of a similar centre-surround antagonistic structure as those in the earlier parts of the visual system. However, some hold more complex structures: *simple* and *complex* cells. Simple cells have an elongated receptive field structure, with flanking antagonistic sub-regions (Hubel & Wiesel, 1959; 1962). This arrangement is believed to be acquired through the spatial alignment of retinal ganglion cells with similar receptive field structures. Thus, these cells rely on the inhibitory processes of earlier mechanisms to detect lines or bars defined by luminance, and are also sensitive to their orientation and spatial frequency. Complex cells, as their name suggests, have an even more sophisticated receptive field structure; as well as being sensitive to the dimensions of orientation and spatial frequency, they have no obvious excitatory/inhibitory sub-regions (Hubel & Wiesel, 1962). Thus, they respond to lines of particular orientations regardless of the exact position at which they are placed within their receptive field. They also show insensitivity to the stimulus' contrast polarity. That is, they are indifferent to a black bar presented on a white background and a white bar on a black background.

Such suppressive mechanisms are achieved through local, lateral inhibition of cellular activity, and have been documented extensively (Macknik & Martinez-Conde, 2004). For instance, it is known from initial work on the limulus horseshoe crab that in a circuit of connected neurons, adjacent neurons have the ability to suppress the activity of one another by means of their own excitation. Many of the core characteristics of the limulus's visual system are found in that of primates, and many aspects of human perception have been successfully accounted for by principles that have arisen from experiments on the compound eye of the limulus. Reports, for instance, have typically suggested that even at the level of the primary visual cortex, uniform luminance exerts little effect on the region's neural activity in comparison to luminance contrast (Dai & Wang, 2011), and instead evokes intracortical suppression (Tucker & Fitzpatrick, 2006). In turn, it is predicted from a human model based on the assumptions of lateral inhibition that the strongest (excitatory) neural activity is found for regions of space immediately inside the borders of an object and the strongest inhibition immediately outside. One observable perceptual consequence of this is the effect found in the Mach band illusion (Ratliff, 1965; Troncoso, Macknik & Martinez-Conde, 2011). In one variant of this illusion, a series of increasingly bright strips are placed side by side. Despite each strip having a uniform luminance, each one appears darker at the side immediately adjacent to the lighter strip, and by implication, lighter at the opposite side. This perceptual quirk can be explained by the effects of lateral inhibition and an important consequence is that the contrast across the boundary is *enhanced,* or a*ccentuated.* This preferential sensitivity to local contrast is echoed throughout the visual system; in primary visual cortex, for example, neurons sensitive to luminance contrast far outnumber those sensitive to uniform luminance (Dai & Wang, 2011). Instead, uniform luminance may evoke intracortical suppression (Tucker & Fitzpatrick, 2006), thus reducing the overall sensitivity to diffuse illumination. The reason why the visual system would adopt separate mechanisms to process luminance and contrast are not easy to understand. We could infer, however, that because in natural scenes luminance and contrast co-exist but are not statistically related (Mante, Bronin, Frazor, Geisler & Carandini, 2005), that perhaps the visual system merely reflects the statistical properties of the stimuli to which it is most sensitive (Dai & Wang, 2011).

1.1.4. Filling in the blanks – how boundaries shape perception

Thus there seems to be little room for the neural representation of surface features in the early parts of the visual system: it is structured such to prioritise the detection of local contrast over regional uniformity. This is strikingly reinforced by a simple psychophysical experiment: a phase-reversing masking paradigm employed by Rogers-Ramachandran and Ramachandran (1998) has shown that contours between regions of light and dark patches are perceived prior to the regions themselves. That is, when the phase (luminance) of the regions is reversed above a particular frequency (~7-8 Hz) it is no longer possible to distinguish the dark region from the light region at any one time, yet a contour is still clearly visible between the two. This is known as flicker-defined form; see **figure 1.1**. Such results imply separate mechanisms for fast contour processing and slow surface perception, a notion which is further supported by the findings from meta- and para-contrast experiments by Breitmeyer, Kafaligönül, Ögmen, Mardon, Todd and Ziegler (2006). So how is regional information ultimately represented by the brain?

It is only with careful experimental manipulation of stimuli that we find that regional information is “reconstructed” to a large extent at a later processing stage on the basis of information gleaned from edge-processing. Such theoretical advances initially stemmed from observing how the visual system overcomes the “blind” regions in the visual field. For instance, in each eye, a natural “blind spot” is present, corresponding to the location on the retina of the optic disc. This is where the fibres of retinal ganglion cells converge to form the optic nerve and exit the retina, thus no photoreceptive cells are present in this surprisingly large retinal area. Despite this, and even under monocular conditions in which the visual system cannot substitute information from the contralateral eye, we are rarely, if ever, aware of a blind spot at all, as the visual system proceeds to “fill in” the region with surrounding information. In addition, if images are stabilised on the retina, similar effects can be observed. For instance, if an annulus of one colour is placed surrounding a disc of another colour and the border between the regions is stabilised on the retina (thus causing neural adaptation and the consequent loss in strength of the border signal) then the colour from the external region appears to “bleed into” or “fill in” the internal region (Krauskopf, 1963; Larimer & Piantanida, 1988). Similar “filling in” effects are found in pathological cases of scotomas (Gerrits & Timmerman, 1969), yet it is perhaps in the former cases that it is most striking, as these demonstrate that such processes are normal aspects of visual processing.

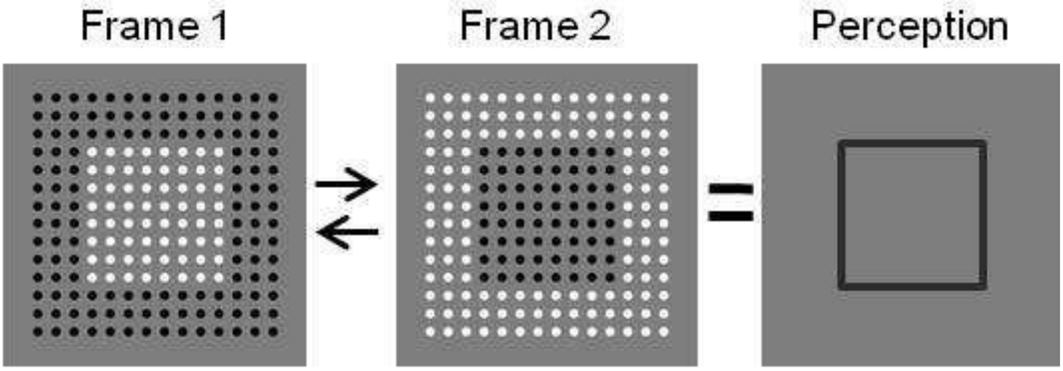

Figure 1.1 Example of flicker-defined form. When the two frames shown above are presented in high frequency alternation, an implicit border can be seen between the inner and outer regions despite it being impossible to tell at any given time which is dark and which is light. Thus, when an intervening black border is placed between the two regions, the viewer cannot perceive a difference between the two regions at such high frequencies. Note that the perception shown in the figure is just an illustration, black lines are not seen.

Indeed, through using techniques of visual masking, some experimenters have highlighted how the perception of surface features is prevented or changed somehow as a result of a subsequent mask. Paradiso and Nakayama (1991), for instance, presented subjects with a central white disc (on a black background) onscreen for 16 ms, followed by a white annulus, smaller than the disc, for 16 ms. When the inter-stimulus interval was sufficiently short, observers did not perceive the central whiteness of the disc, instead believing it to be dark (like the background; see **figure 1.2**). This was explained in terms of a border-to-interior filling-in process, by which the perception of the central contour of the annulus prevented the successful completion of the filling-in of the initial disc, which proceeded from the outer edge (corresponding to the location of contrast).

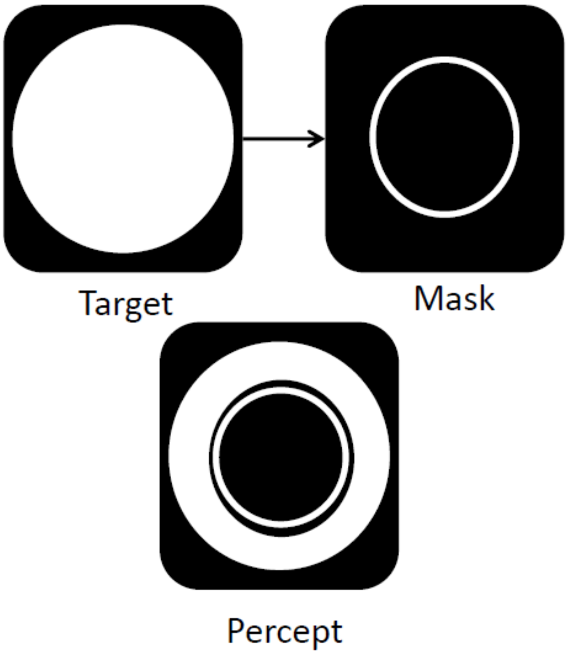

Figure 1.2 A cartoon illustration of the filling-in processes revealed through backward-masking (adapted from Paradiso & Nakayama, 1991). When the white disc (target) is presented for a brief period (16ms) and followed immediately by the smaller white annulus (mask), the resultant perception is that there is a “hole” in the centre of the white disc, corresponding to the location of the mask. This is explained by a theory in which the visual information is first represented by the visual system from the contours present in the image (i.e. first from the edge of the disc before proceeding inwards). The presentation of the mask’s contours within the disc impedes this process and so the neural representation of the regional information within the disc is not completed.

The perceptual effects of representing regional information subsequently to, and dependently on, edge information are demonstrated very well in a simultaneous lightness contrast illusion, in which two regions of equal luminance are embedded in two different backgrounds, one lighter and one darker. The resultant perception is that these regions can look very dissimilar, with their relative luminance being shifted in the opposite direction to that of their surrounding backgrounds (Diamond, 1953; Heinemann, 1955; see **figure 1.3a**). Remarkably similar effects are also observed in instances of simultaneous colour contrast, in which the stimuli vary in chromaticity, not luminance (Kingdom, 1997). That is, the appearance of surface colour to an observer is heavily dependent on the chromatic contrast present at the border between figure and ground, such that the perception of the colour of an interior region is shifted towards the opponent colour of a surrounding region (see **figure 1.3b**). Such illusions are striking demonstrations of just how little the perception of surface features correlate with the actual information that is entering the eye, and in contrast how much they are influenced by information found at adjacent boundaries.

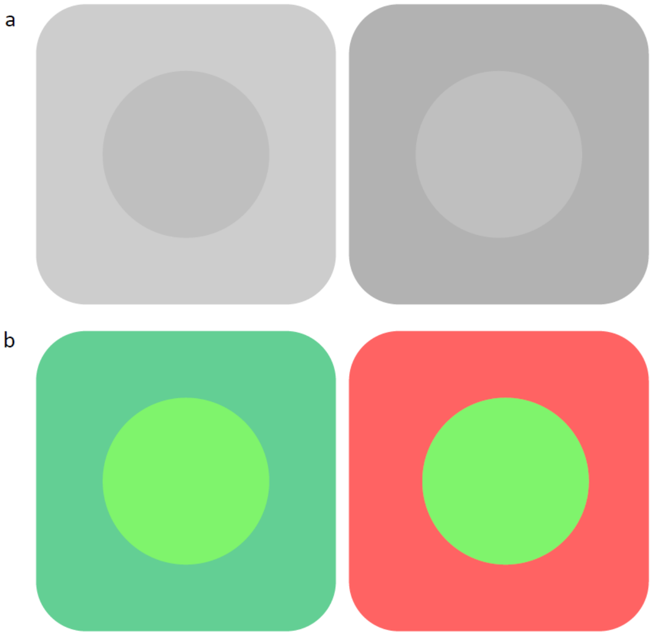

Figure 1.3 Demonstrations of simultaneous brightness (**a**) and colour (**b**) contrast. In both rows, the left and right circles are identical. The presence of the surround, however, can significantly shift the perception of the circles' surface brightness/colour in the opposite direction to that of the surround relative to the brightness/colour of the circle. As a result, the circle on the top right appears brighter than that on the top left, and the circle on the bottom right appears "greener" than that on the bottom left.

Why does the visual system rely so heavily on what is present at the border, and not within the region itself? One answer is that by encoding simultaneous chromatic (and luminance) contrast, the visual system is able to achieve some level of perceptual constancy (Ekroll, Faul & Niederée, 2004; Hurlbert & Wolf, 2004). What is pervasive in colour perception, for example, is the challenge faced by the visual system in perceiving a *constant* surface colour despite changes in the spectral composition of the light reflecting from its surface, which arise from natural changes in illumination. It is important that the visual system attributes the change in visual information to the environmental conditions rather than, incorrectly, the object colour in order to maintain a consistent representation of the world. One method of achieving this is to directly encode colour contrast, as the ratio in cone excitations between surfaces tends to be constant through changing illumination. It is only through the use of ambiguous, experimental stimuli in which the shortcomings of such a mechanism can become apparent, as with the simultaneous contrast illusion described above, which lead to a gross misperception in surface colour. Importantly, however, in both cases of simultaneous luminance and chromatic contrast, surrounding each inner region with a black border, thus separating each region from its background, diminishes the effect (see **figure 1.4**). This illustrates that it is the presence of a *contrast* (i.e. an informative border) between the figure and background that mediates the effect.

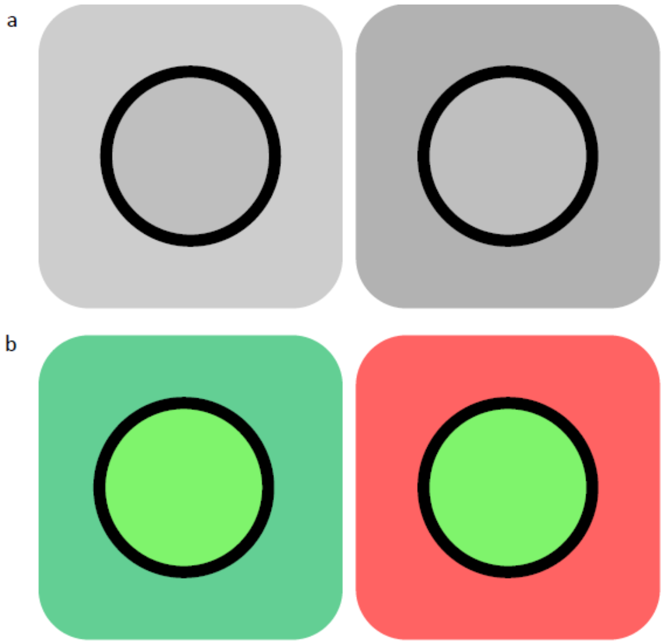

Figure 1.4 A reproduction of the illustration shown in **figure 1.3**, but with added black borders. These borders mask the contrast between the circles and their surrounds, thus diminishing or completely abolishing the simultaneous contrast illusion. This demonstrates that the information that is gleaned from edge-based processes can significantly shape the way that regional information is perceived.

This is also reflected in the physiological structure of the mammalian visual system, as revealed through neuropsychology and cell-recordings. In colour perception, for example, cortical damage to ventromedial areas can result in a complete and highly selective loss of colour experience, known as cerebral achromatopsia (Heywood, Cowey & Newcombe, 1991). In spite of this profound deficit, however, one patient in particular has been shown to remain sensitive to chromatic contours (Heywood, Cowey & Newcombe, 1994). The neural locus of this spared ability is likely a lot earlier than that of his impaired abilities. It is presumably the primary visual cortex, which is known to signal local cone-contrast through its “double-opponent” cells (Kentridge, Heywood & Cowey, 2004). Indeed, signals arising from these neurons are necessary in the mechanisms of colour constancy (as previously discussed), as without these neurons (through selective damage to primary visual cortex) the illusion of simultaneous chromatic contrast fails (Kentridge, Heywood & Weiskrantz, 2007). The implication from these findings is that surface colour is reconstructed at a later cortical level on the basis of signals arising from earlier neurons sensitive to chromatic contrast at borders.

Importantly, however, such border-to-interior “filling in” processes are not restricted to luminance and colour perception. Regional features involving textural elements (2nd-order stimuli) are also perceived on the basis of some filling in mechanism (Ramachandran & Gregory, 1991). Caputo (1998) and Motoyoshi (1999), for instance, adopted the same principal method as Paradiso and Nakayama (1991) but with oriented line textures. In line with the original findings, Motoyoshi found that the perception of the line segments of the target disc was suppressed inside the area of the annulus mask. This is completely in line with the perception of the interior region of the disc spreading from the outer edge, which is subsequently interrupted by the presence of the contour in the mask, thus preventing complete filling in. Additionally, however, if the central region of the disc contained some regional discontinuity (i.e. mis-oriented line segments) then the perception of this would survive the effect of the mask (see **figure 1.5**). This is strong evidence that regions of local discontinuity, even with texture stimuli, gain precedence in the temporal sequence of object processing by the visual system. Recently, Su, He & Ooi (2011) found that by manipulating the stimulus onset times of a grating textured surface, they could demonstrate that the perception of that surface begins at the location of borders and proceeds inwards, resembling the classic border-to-interior processing strategy. In general, the results suggest that regardless of how a shape may be defined, it is always the perception of local contrast between figure and ground that leads.

Neuroscientific evidence corroborates with these psychophysical findings. Lamme, Rodriguez-Rodriguez and Spekreijse (1999), for instance, found single-cell evidence in the awake macaque of a response enhancement at the location of an orientation contrast that occurred prior to any location within the regions of the textures themselves. The largest response latencies were recorded for regions farthest away from the contours. Similar findings were recorded by Romani, Caputo, Callieco, Schintone and Cosi (1999) using visual evoked potentials, which, together, corroborate with Motoyoshi (1999) in suggesting the workings of a texture analysis mechanism that prioritises locations of contrast relative to regions of uniform texture. On average, the surface of a textured region might be filled in between 40 and 60 ms following boundary localisation. Thus, it is believed that once the contours within an image are defined by the visual system, the interior regions of those contours are then filled in, a process which relies on re-entrant feedback from higher visual areas. The perceptual process of texture "filling in", however, is likely to differ from that of regular "luminance" or "colour" filling in, and may involve a two-step process (Caputo, 1998). Initially, the average luminance intensity of a region may be established and "spread" within the previously detected boundaries (lasting between 40 and 80 ms), following which the filling-in of the individual textural elements occurs (i.e. their shape, orientation, size).

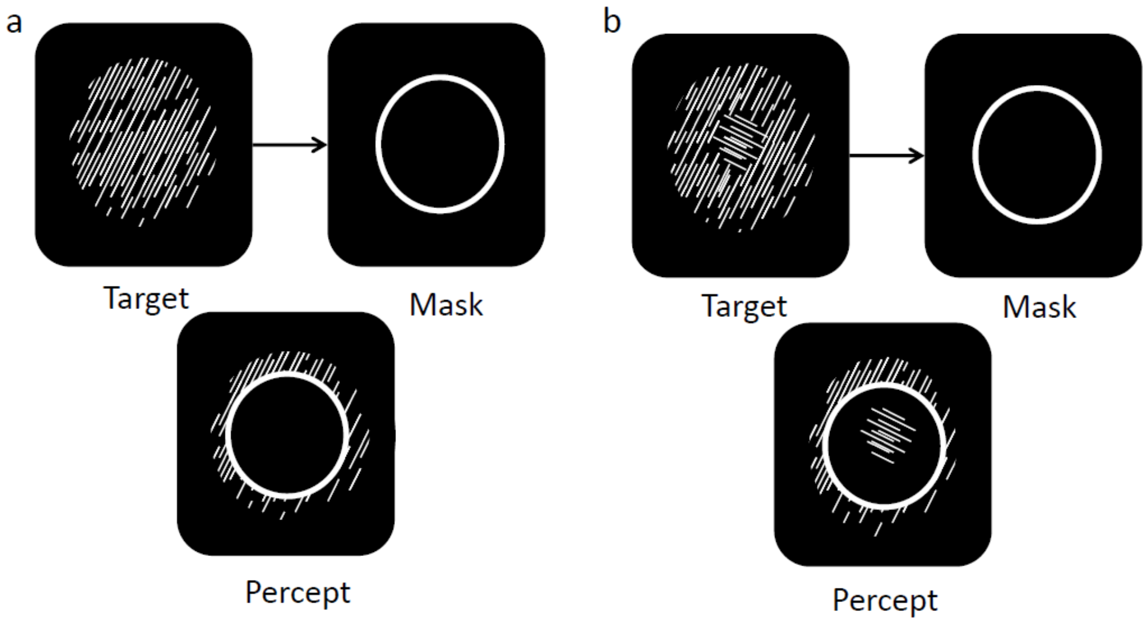

Figure 1.5 A cartoon illustration of the texture-based filling processes revealed through backward masking (adapted from Motoyoshi, 1999). **a** shows an effect that is analogous to that shown in **figure 1.2**, in which the perception of a textured region is disrupted by the mask's prevention of the filling-in processes. **b** shows that this effect does not entirely uphold when a region of contrast is present inside the textured region. Because the filling-in process begins at locations of contrast, the perception of the sub-region persists despite the mask.

The influence of borders of contrast is also present for other regional features. Lateral inhibitory mechanisms also underlie the visual system's processing of spatial frequency (Sagi & Hochstein, 1985), orientation (1-D lines, Blakemore, Carpenter & Georgeson, 1970; 2-D plaids, Smith, Wenderoth & van der Zwan, 2001) and image contrast (Chubb, Sperling & Solomon, 1989; see **figure 1.6**). For example, for a circular field of vertical lines, a surrounding annulus consisting of lines tilted clockwise will "repel" the apparent orientation of the central lines in a counterclockwise direction (see **figure 1.7**). This attribute of the visual system is pervasive and underlies much visual functioning, again reinforcing for us the notion that the perception of regional information is second to what is perceived at local areas of contrast. The visual system's relatively high sensitivity to discontinuity in an image reflects the assumption that most objects in a visual scene have a large amount of surface uniformity with few (if any) abrupt discontinuities, rendering a large portion of the retinal input redundant in terms of its importance in (early) visual processing (Chong & Treisman, 2003). The primacy of boundaries over surface information is reflected in findings from psychophysics and behavioural experiments showing that information about form is integrated outside the workings of awareness (Fahrenfort, Scholte & Lamme, 2007; Breitmeyer, Ogmen, Ramon & Chen, 2005) and prior to the integration with surface features (Biederman & Ju, 1988; Elder & Zucker, 1998; Humphreys, Cinel, Wolfe, Olson & Klempen, 2000).

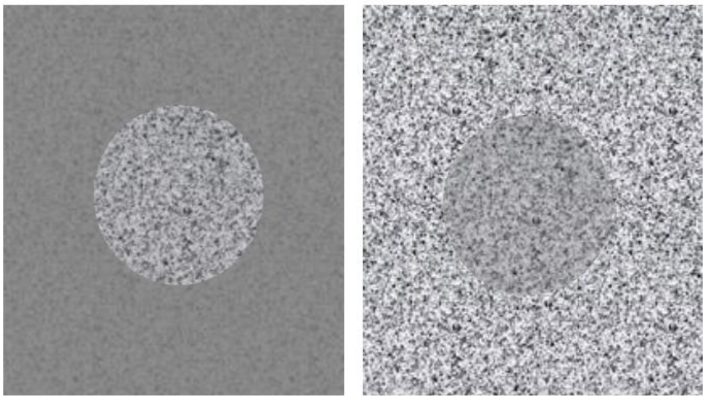

Figure 1.6 The Chubb contrast illusion (Chubb *et al,* 1989). The two inner circular regions are identical; the one on the left, however, is surrounded by a lower contrast background and so appears to be of a higher contrast than that on the right.

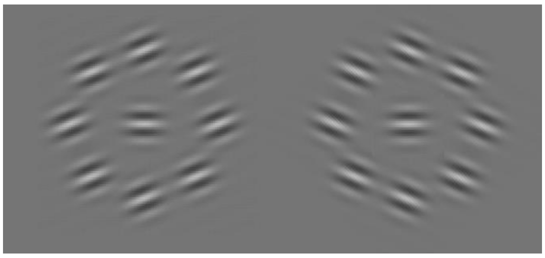

Figure 1.7 Demonstration of the simultaneous tilt-illusion. Both central Gabor patches are horizontal, yet the presence of the tilted surrounds (clockwise on the right, counter-clockwise on the left) shifts the perception of the central orientations in the opposite direction. As a result, the central Gabor patch on the left appears tilted clockwise from horizontal, and the one on the right appears tilted counter-clockwise from horizontal.

1.2. Mechanisms of 2nd-order segmentation and discrimination

1.2.1. Filter-rectify-filter

Natural scenes typically contain contours associated with discontinuities in object surfaces, or which separate figures from their ground (Anzai, Peng & van Essen, 2007). Although we may most often associate such contours with spatial changes in colour or luminance, or less obvious cues such as binocular disparity, it is equally conceivable to imagine an object that stands out from its background purely on the basis of its surface texture (or its 2nd-order properties). Thielscher and Neumann (2003) present a compelling example, for instance, in which the outline of a Dalmatian is perfectly visible against a background of a black and white checkerboard pattern. Importantly, both the figure and ground share the same average luminance intensity, only differing in the textural arrangement of their regional elements. Despite this, a clear contour is visible to us at the figure-ground boundary. Such observations are taken as direct evidence by some researchers that the visual system has access to higher order properties of surfaces that allow for a preattentive contrast to be formed between a figure and its ground purely on the basis of texture (e.g. Bergen, 1991; Palmer & Rock, 1994; Landy & Graham, 2004). Although texture-based segmentation is a robust and widely studied attribute of the visual system, no clear consensus exists as to what distinguishes one texture from another to allow this effortless segmentation.

Julesz' work, however, is perhaps the most widely read and cited in this field, as he was among the first to attempt to establish a foundation for the understanding of the components of visual texture (Julesz, 1962). Although Julesz' original efforts centred around the role that certain visual statistics played in the automatic segmentation of patches of randomly distributed luminance dots (Julesz, Gilbert, Shepp & Frisch, 1973), counterexamples to these theories, discovered by Julesz himself, made the evidence difficult to interpret (Caeli & Julesz, 1978; Julesz, Gilbert & Victor, 1978). These counterexamples were suggestive of an alternative explanation, and consequently Julesz' work became synonymous with the "texton", an easily discernible image feature (i.e. a corner, termination etc.) that allows texture patterns to segment despite identical pattern statistics (Caelli, Julesz & Gilbert, 1978). Such a process in the visual system is attributable to the early retinal and cortical 'feature extractors', which is a conception not too dissimilar to the findings of Gestalt psychologists such as Beck (1972, 1982), who concluded after many experiments that segregation resulted from the feature density of sample attributes, such as luminance, colour, size, orientation and movement. These conclusions were met with scepticism by some researchers, however, who claimed that the image features/textons were not well enough defined.

Consequently, an alternative line of modelling arose, which concerned the orientation and spatial frequency tuned channels outlined in spatial vision research (De Valois & De Valois, 1988; Graham, 1989, 1992). Knutsson and Granlund (1983) theorised that the distribution of power in different spatial frequency bands might be used to segregate natural textures from one another. Strong evidence for such a model came from the findings of Bergen and Adelson (1988), which showed that two patches of segmenting textures could be made less conspicuous by increasing the size of the components in one patch to equate the dominant spatial frequency of the different textural elements. This is evidence against a simple texton theory, and has led to the evolution of a number of computational models that attempt to explain how an ideal system would compute a texture gradient (a spatial difference in textural composition, i.e. a contour). A simple luminance edge, for example, can be defined by a band-pass linear spatial filter similar to a cortical simple cell, which would produce a peak response at the boundary of contrast. A typical texture-defined edge, such as Thielscher and Nuemann's (2003) Dalmatian, however, has the same average luminance on either side and so can not be detected by a purely linear mechanism. Many computational theories were constructed circa 1990 on the basis of known cortical cell types, and are so similar in nature to one another that they can be grouped together under the collective title the "back pocket model" (Bovik, Clark and

Geisler 1990; Caelli 1985; Fogel and Sagi 1989; Graham; 1991; Landy and Bergen 1991; Malik and Perona 1990; Sutter, Beck and Graham, 1989; Turner, 1986). The basic formula of such a model consists of a set of linear spatial filters, which model the output of V1 cells, an intermediate non-linear transformation stage containing lateral inhibition to suppress spurious weak responses and a final process of linear filtering used to enhance texture borders.

Imagine two abutting patches of uniformly distributed orientation signals, otherwise identical to one another except that one consists entirely of vertical signals, the other of horizontal signals. They sit immediately adjacent to one another, forming a texturally defined vertical edge. When applied to the whole of the retinal image, a spatial filter that is sensitive to a vertical orientation, for example, will elicit a weak response across the entire region defined by horizontal signals. Across the vertically-defined texture, however, the response will be intensely variable due to very strong positive signals from the vertical elements and very negative signals from regions just to the side of each element. At this stage, the average neural output for each patch is comparable to the other until a rectifying nonlinearity is applied to the whole image, resulting in no change to the signals representing the horizontal region, but a much larger response in the signals representing the vertical region. Finally, the output of the previous rectification is in turn analysed with yet larger scale spatial filters, with a vertical filter yielding the strongest response to signal the presence and location of the vertical texture border. This computational modelling approach, and the work of Julesz (1981), Beck (1983) and Bergen and Adelson (1988), has given strong support for the importance of basic image attributes, especially orientation, in automatic texture segmentation, whereas elements such as “crossings” or “terminations” have been shown to have some contribution, albeit much weaker. The term 2nd-order segmentation thus describes the edge-based processes that detect discontinuities in visual attributes that can not otherwise be picked up by 1st-order linear mechanisms.

1.2.2. Cortical contributions to 2nd-order segmentation

Despite dedicated efforts to do so, it has been difficult to isolate a cortical region in the visual system that can be said with much confidence to perform texture segmentation. This is no more obvious than in the work on single-cell recordings in animals and the results from human neuroimaging, which together often lead to contradicting results. Additionally, the notable lack of neuropsychological patients with a complete and unique impairment in texture segmentation suggests that we may be considering a process that is achieved through recurrent activity between many parts of the visual system (see Bergen, 1991). Computational models based on the anatomical and functional properties of the areas of the visual system are also supportive of this theory (see Thielscher & Neumann, 2005).

The earliest possible stage of visual processing in the cortex, V1, is seen as a worthy substrate for at least the linear filtering stage of texture segmentation. This is due in part to its extensive network of simple and complex cells, and the presence of both excitatory and inhibitory lateral interactions. Such complexity is speculated by some researchers to be sufficient to perform even 2nd-order nonlinearity (Landy & Graham, 2004). Consequently, V1 is a strong candidate for a role in enhancing responses to borders of pop-out stimuli (Kastner, Nothdurft & Pigarev, 1997, 1999; Nothdurft, Gallant and van Essen, 1999). The first hints that V1 could potentially conduct textural analyses came from single-cell recordings in the macaque, which highlighted that the activity of V1 cells in response to an oriented line or texture was modulated by the presence of an orthogonal surround (Knierim & van Essen, 1992; Sillito, Grieve, Jones, Cudeiro & Davis, 1995). Importantly, some researchers have highlighted the independence of V1 in this respect: Lamme and colleagues (Lamme, Supèr, & Spekreijse, 1998; Lamme *et al,* 1999), for instance, showed that recordings from cells sensitive to the border of a shape defined by an orientation contrast of 90° did not diminish with substantial ablation of extrastriate cortex, whilst Hupé, James, Girard and Bullier (2001) reported similar effects for responses found in macaque V1 to a surrounding texture when V2 was suppressed. Furthermore, Nothdurft, Gallant and van Essen (2000) showed V1's neural activity to be enhanced at the boundary but suppressed at regions of homogenous texture in the anaesthetised macaque. Such results indicate a role of the primary visual cortex in signalling the presence of a textural boundary, with some authors citing the horizontal long-range connections within this area as good mediators of the observed border effects, and potentially in isolating figure from ground or detecting surface discontinuities in simple image attributes (Thielscher & Neumann, 2003).

The primary visual cortex, however, is not unique in its apparent sensitivity to texture; its immediately adjacent topographical cortical area (V2) has recently been shown to contain a proportion of neurons that show inhibitory interactions, making them ideal candidates for the signalling of orientation combinations (Anzai, Peng & van Essen, 2007), a process which may play a pivotal role in combining and contrasting spatially separated texture information. V2 cells, both at the individual and population level, also respond well to complex shape information related to object texture and contour information (Hegdé & Van Essen, 2000, 2003) and in a way that evolves relatively quickly (as early as 40ms of stimulus exposure; Hegdé & Van Essen, 2004). In a direct comparison between the activity of V2 and V1 neurons, von der Heydt and Peterhans (1989) found a significant response to texture contour stimuli in V2, but not in V1. Ramsden, Hung and Roe (2001) also provided similar evidence. Together, the response properties of V2 cells suggest vital processes in the general recognition of form by complex shape. This is largely upheld by the finding that lesions of V2 produce significant impairments in segmenting a texture-defined figure (Merigan, Nealey & Maunsell, 1993). That is, the orientation of a figure comprising a subset of line segments of different orientations to their background could not be discriminated, whereas the same figure defined by colour or size could be.

The role of higher visual areas in similar tasks is less clear. Merigan (2000) lesioned area V4 in monkeys and demonstrated that the ability to segment two abutting regions of orientation orthogonal to one another was not abolished. Counter-intuitively, however, the same animals could not discriminate the orientation of a bar composed of the same orientation contrast as in the previous task. This may simply have been a consequence of the difference in task difficulty, or may have highlighted a dissociation between the low-level process of segmenting, which remains intact, and the high-level process of figure perception, which is abolished. Evidence from simulation studies based on computational modelling has contributed to our understanding of the role of individual visual areas in texture segmentation. These models consist of mock areas V1, V2 and V4 as well as their interconnections, and are formulated on the basis of data gleaned from both behavioural and neural work (see Thielscher and Neumann, 2005, & 2007). The visual areas within the model each hold a topographic map of cells selectively tuned to orientations. By removing the feedback connections in the model (i.e. creating artificial "lesions"), it becomes possible to understand the contribution of individual visual areas, and their feedback modulations, to the segmentation of texture information. One of the main findings is that in a full recurrent-flowing model, a texture bar defined by the offset of oriented line segments from their background is represented moderately by V2 but very strongly by area V4. When feedback from V4 is removed, however, the observed stimulus-specific activity in V2 notably diminishes, and it can no longer signal the presence of a textural boundary. The role of feedback activity, at least within these models, is thus essential to segmentation.

Neuroimaging evidence with humans also supports the role of later visual areas (in particular V4) in the process of texture segmentation. For instance, Thielscher, Kolle, Neumann, Spitzer & Gron (2008) varied the orientation contrast of a figure defined by a set of oriented line segments in a field of similarly oriented background elements. The preattentive BOLD response in early visual areas (V1 and V2) was not correlated with the orientation contrast, whereas the response in higher areas (V4v, VP and lateral occipital complex (LOC)) was. This corroborates somewhat with previous findings (Kastner, de Weerd and Ungerleider, 2000; Schira, Fahle, Donner, Kraft & Brandt, 2004; Larsson, Landy and Heeger, 2006) although Thielscher *et al* (2008) were the first to show the correlation with behavioural psychophysical data, thus supporting a trend between observed neural activity and perception. There is also evidence showing these higher areas to have greater stimulus-specific adaptation than lower areas when participants viewed patterns containing a border defined by the offset of two sets of oriented lines (Montaser-Kouhsari, Landy, Heeger & Larsson, 2007). Together, these results indicate that the final output of the segmentation is primarily represented in the mid level areas (esp. V4), and any observed activity in relatively low-level areas may simply reflect modulatory feedback. It is not surprising, then, that patient HJA, with lesions to the human equivalent of V2, V3 and V4, is unable to segment a texture-defined bar (Allen, Humphreys, Colin & Neumann, 2009). In corroboration with the neuroimaging evidence, it seems likely that it is damage to the mid-level area V4 that is crucial to his impairment.

The neuroimaging data, however, are not entirely convincing. Although a lot of the evidence indicates V4 as the crucial mediator in texture segmentation, its activity is only a correlation with what is being perceived. The LOC, on the other hand, has been extensively studied in separate investigations and seems to represent a neural correlate of the perception of object shape seemingly regardless of how the shape is defined (e.g. Kourtzi & Kanwisher, 2001; Tyler, Likova & Wade, 2004). The activity observed in this region, therefore, may not directly reflect the segmentation process *per se*, but that of a general cue-invariant integration stage beyond the early- to mid-cortical processing hierarchy (Thielscher *et al,* 2008). The reason why its activation is not observed in some of the neuroimaging evidence (Kastner *et al,* 2000; Schira *et al*, 2004), but it is observed in others

(Thielscher *et al,* 2008; Larsson *et al*, 2006) may instead be due to differences in task demands and relate more generally to object recognition processes.

Some authors have even argued that the stimulus-specific activity in V4 in response to a texturally defined figure (Thielscher *et al,* 2008; Kastner *et al,* 2000; Schira *et al,* 2004; Larsson *et al,* 2006) can be explained by an indirect result of feedforward processing from V1 and V2. Hallum, Landy and Heeger (2011), for example, provided evidence from fMRI-adaptation that shows populations of neurons within V1 to be sensitive to 2nd-order contours. In addition, in Schira and colleagues' (2004) study, V2 (although not V1) was shown to have systematic changes in BOLD activity in response to the degree of texture contrast that defined a figure. In this study, subjects were required to complete an attentionally-demanding foveal letter-discrimination, thus preventing the peripherally presented texture contours from entering awareness. This is important, as the observed activity is thus likely to reflect that which is independent to attention-related components of contour perception; the activity is related to the pre-attentive (and pre-awareness) component of segmentation. Although Kastner *et al* (2000) manipulated attention and did not find early cortical activity related with texture-defined figure perception, they did not dissociate activity related to the figure *per se* and that of the contour. Schira *et al* (2004) did, however, which is perhaps why they found significant stimulus-related activity within V2.

In conclusion, it is currently unknown exactly where in the cortical processing hierarchy texture segmentation is achieved. One theory is that a recurrent flow of activity is required between areas V1-V4, with the final neural representation of the texture figure occurring in V4. Another quite different theory, however, is that the early cortical stages (specifically V1 and V2) are sufficient, and the activity found in V4 merely reflects feedforward projections from these areas.

1.2.3. Regional discrimination of 2nd-order stimuli

Texture segmentation is not the only, and perhaps not even the primary, example of the use of 2nd-order stimulus information by the visual system; Landy and Graham (2004) proposed three motivations for the visual system's computation of visual texture, only one of which concerned the segmentation of one textured surface from another. The remaining two, namely 3D structure perception and object recognition, differ from the previously discussed motivations for the use of texture as they do not rely on the immediate segmentation of one textured patch from another. After all, we are capable of discriminating patches of isolated texture and describing them in great detail. For the visual system, it would be easy to imagine how this would pose much more of a complex task than that of segmentation. Such a computation requires the integration of spatially distributed local signals, which can be informative in reliably estimating visual texture (Dakin & Watt, 1997; Dakin, 2001). When applied to the perception of objects, this process is potentially vital in determining their surfaces under varying viewing conditions (Dakin & Watt, 1997; Bovik, Clark & Geisler, 1990; Victor & Conte, 2004). In particular, extracting the orientation statistics of distributed elements has often been considered as a potential foundation for the discrimination of certain surface textures (Witkin, 1981; Kass & Witkin, 1985; Blake & Marinos, 1990; Field, Hayes & Hess, 1993; Dakin & Watt, 1997; Dakin, 2001; Wolfson & Landy, 1998).

This component of visual texture analysis, “orientation pooling”, has been studied in some depth by some researchers. Dakin and Watt (1997), for instance, showed that when confronted with a spatially unstructured texture, observers made accurate judgments of the mean orientation of globally distributed elements, as well as the variance of the distribution from which the elements were drawn. Dakin (2001) has since shown that, by assessing the effects of a texture’s size, density and numerosity on the observer’s internal noise and the sampling density, observers relied on a sample size approximately equal to a fixed power of the number of samples present. This flexibility in orientation pooling is thought to be a prerequisite for efficient texture perception, in which accurate estimation of the texture’s appearance requires the integration of spatially distributed elements and the extraction of some statistical quality that does not vary with changes in viewing conditions. Neuropsychological evidence from a human patient has also indicated the role of V4 in tasks of a similar description: patient HJA, as discussed before, is unable to correctly scale the sampling size of his estimate when confronted with an unstructured patch of orientation signals with increasing signal noise, which suggests an inability to extract a reliable statistical representation and thus an inability to perceive the spatial relationship of distributed signals (Allen, Humphreys & Bridge, 2007). As mentioned previously, any conclusions drawn from this patient must be used tentatively as his lesion covers areas V2 and V3 as well as V4. On a related note, Beason-Held and colleagues (1998) have shown that when observers were asked to view random-dot texture patterns and spatially correlated patterns of the same stimuli, activity in the primary visual cortex, as well as the middle lingual and fusiform gyri, was recorded for random patterns, but for spatially correlated textures, activity extended into middle temporal areas. This suggests a selective role of potentially higher visual areas in computing the spatial relationships of distributed visual signals.

From a different approach, Wolfson and Landy (1998) presented observers with two adjacent patches of texture defined by the distribution of individual oriented elements and asked them to detect a difference. If the patches differed in mean orientation, then detecting a difference was easier if the patches were abutting than if they were separated. If the patches differed in the variance of the distribution from which the orientations of their respective elements were drawn, however, then no such effect was observed. This is indicative of a higher-order image statistic (i.e. orientation variance) being associated with a region-based mechanism; in other words not related to texture segmentation, but more related to the appearance of a surface. Caution should be taken, however, in interpreting this evidence as there was a confound of an increased quantity of line crossings in patches with greater orientation variability. This renders a direct interpretation of “edge-based” vs. “region-based” mechanisms dubious as the observer could essentially use a separate cue in the discrimination-of-variation task to that used (i.e. orientation sampling) in the discrimination-of-mean task. Whether or not this factor accounts for the observed effects of discriminating patches of texture that are abutting and separate must be explored before reliable conclusions can be drawn. Potentially, if the conclusions from Wolfson and Landy (1998) are accurate they may indicate a distinction within the visual processing of texture depending on its intending function: edge-based texture mechanisms are most useful in segregating an object from its background, whereas region-based mechanisms are most informative in deducing the characteristics of a texture as a property of an object’s surface. Explorations into the neural substrates of these separable processes are yet to be made.

1.2.4. Representation of texture as a physical surface property

Determining the spatial characteristics of distributed textural information is likely to subserve high-level object recognition. The appearance of the geometry of an object’s surface, for instance, requires the integration of spatially distributed cues, and it is evident from some of the evidence discussed thus far that, rather than comparing by exemplar on a pixel-by-pixel basis, the visual system must at least extract a statistical feature that is common across viewing conditions in order to achieve an accurate perceptual representation of the characteristics of a textured surface (Cho, Yang & Hallett, 2000; Victor & Conte, 2004). This is a vital process in achieving perceptual constancy when describing the structure of objects. What are the processes that determine a golf ball's characteristic surface? For the visual system, this is not an easy question to answer; a golf ball's surface is something that is highly influenced by changes in illumination and the local height variations of the textural elements, as well as the viewing geometry (Nishida & Shinya, 1998; Fleming, Dror & Adelson, 2003; Pont & te Pas, 2006). This ensures that individual pixel values acquired from a visual image vary considerably despite no changes in the physical structure of that image. Despite these limitations, we are able to describe an object's structure at the levels of mega-, meso- and micro-scale (Koenderink & van Doorn, 1996; Ho, Landy & Maloney, 2008). A ping-pong ball, for example, holds a similar megastructure to a golf ball; both are spherical. Their respective meso-structures, however, are radically different; the ping-pong ball's surface is smooth, the golf ball's furrowed. The molecular organization of the two objects' material also denotes their microstructures, which produces a characteristic glossiness on the golf ball that is absent on the ping-pong ball. It is the latter two which are potentially good indicators of an object's material, i.e. the "stuff" of the visual world, whereas the first is most useful in outlining the forms of objects, i.e. the "things" (Adelson & Bergen, 1991). In terms of our understanding of the visual system's ability to process objects at these levels, we know very little about the processing of "stuff" compared to that of "things".

How exactly these computations are performed by the visual system, and where in fact they take place, is a far-reaching question that has only limited answers at the moment. Some research within the last 15 years has made notable attempts to tackle the question using a variety of techniques. For instance, evidence from single cell recordings in the macaque revealed a sensitivity of V4 neurons to the classification of natural textures (Arcizet, Jouffrais & Girard, 2008). Specifically, a large majority of neurons responded to patches of natural textures flashed in the receptive field of V4 neurons in the awake macaque and, most importantly, the responsiveness of the neuron to the stimulus did not change as the stimulus was moved such that the receptive field sampled a different region of the same texture. Additionally, the responses were comparable for the same patch of texture viewed under conditions of changing illumination. V4 neurons have also been shown to be selective to the density and size of points mimicking a shaded granular texture surface (Hanazawa and Komatsu, 2001). Together, these results are suggestive of V4 neurons being position-and illuminant-invariant to texture; possibly being sensitive to some statistical quality of a repeated pattern that is constant under different viewing condition. This is an essential process in the perception of texture as a natural and informative property of an object.

With human subjects, Puce, Allison, Asgari, Gore and McCarthy (1996) utilised functional neuroimaging when normal subjects viewed images of faces, letterstrings and textures. The textures were not composed of identifiable individual elements and certainly did not look artificial, as in the traditional line of work (e.g. Bergen & Adelson, 1988), but instead resembled more realistic materials such as sandpaper, wood and grass. Activity in response to these textures was specific to regions of the collateral sulcus. This was strong initial evidence for a functional specialization of an extrastriate region in processing visual texture. Peuskens *et al* (2004) found similar activation in occipito-temporal regions when subjects made judgments about surface texture in 3D objects. The most notable advancements in uncovering the neural correlates of natural texture perception, however, have been made in the last 3-4 years. Cant and colleagues (Cant, Arnott & Goodale, 2009; Cant & Goodale, 2007) used functional magnetic resonance adaptation to demonstrate selectivity in the LOC and collateral sulcus in observers when passively viewing objects that would change in either global structure or physical texture, respectively. Very recently, Cavina-Pratesi, Kentridge, Heywood and Milner (2010a; 2010b) combined functional magnetic resonance adaptation with evidence from visual agnosic patients. They showed in a double dissociation that making judgments about the form of an object preferentially activated the LOC, whereas judgments about the same object’s texture or colour activated the posterior collateral sulcus and anterior collateral sulcus, respectively. Two neuropsychological patients, one with a damaged LOC yet otherwise spared ventral stream, the other with damage to the collateral sulcus with spared LOC, showed the expected relative deficits in accordance with the neuroimaging findings. Taken together, these recent advancements are highly indicative of a major role of mid-level ventral regions in processing texture as a surface property of an object, especially regions that correspond to the collateral sulcus in humans and area V4 in monkeys.

1.3. The direction of this thesis

The first four sets of experiments documented in this thesis are motivated by the work that has been discussed in this general introduction. Specifically, the first set of experiments will directly compare the processes of 2nd-order segmentation and discrimination in a psychophysical demonstration in normal observers. Segmentation is explored using stimuli that are composed of oriented line segments and in which there is a difference in mean orientation between two abutting regions. A novel type of stimulus will be introduced in these experiments, which will demonstrate the large temporal superiority for segmentation (by spatial gradients in mean orientation). In contrast, regional discrimination will be explored using similar stimuli but which differ in orientation variance. Like differences in mean orientation, differences in orientation variance necessitate mechanisms of a higher order to be detected; unlike differences in mean orientation, however, these stimuli do not result in segmentation. Independently manipulating the first- and second-moment statistics of orientation information, the mean and variance respectively, therefore provides a good way of isolating the edge- and region-based components of 2nd-order vision without confounding changes in other attributes.

In the second set of experiments, it will be explored whether a single patient (patient MS) with bilateral medial occipitotemporal damage (including human “V4”) retains an intact ability to perform 2nd-order segmentation. There is some debate that has arisen from neuroimaging and computational modelling studies regarding the necessary role that human “V4” plays in 2nd-order segmentation; the segmentation-specific stimuli introduced in the first set of experiments offers a novel and effective method with which to test this necessity in an ideal neuropsychological candidate. The focus of the third set of experiments will be on the discrimination of orientation variance. Specifically, through the technique of selective adaptation-induced aftereffects, it will be determined whether this second moment orientation statistic is encoded explicitly by the visual system through channels tuned broadly to high and low levels, and whether this is represented at a stage that encodes the statistic independently of the low-level statistic (mean orientation). Additionally, it will be explored whether this encoding is limited to a retinotopic reference frame, or whether it is achieved through the spatiotopic encoding of high-level cortical areas, and whether patient MS's intact early visual areas are sufficient to allow some sensitivity to the property of orientation variance.

The fourth set of experiments will consider the role of extrastriate areas in the temporal integration of different 2nd-order cues for the purposes of cue-invariant segmentation. Work involving psychophysical techniques in normal observers as well as neuropsychological work with a patient with bilateral damage to LOC will be presented.

The remaining part of this thesis will consider a cognitively higher-level aspect of perception that stems from segmentation. Specifically, the fifth and final set of experiments explores whether objects that are rapidly extracted by the visual system's 2nd-order channels are capable of being selected as units of attention without ever reaching the level of awareness. The relationship between attention and awareness will be reviewed prior to presenting the experiments of study five.

2. Study One - Contrasting 2nd-order segmentation and discrimination

This chapter has been published in *Vision Research*. The full citation is:

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; doi: 10.1016/j.visres.2011.07.021

2.1. Abstract

Regions of visual texture can be automatically segregated from one another when they abut but also discriminated from one another if they are separated in space or time. A difference in *mean orientation* between two textures serves to facilitate their segmentation, whereas a difference in *orientation variance* does not. The present study further supports this notion, by replicating the findings of Wolfson and Landy (1998) in showing that judgments (odd-one-out) made for textures that differ in mean orientation were more accurate (and more rapid) when the textures were abutting than when separated, whereas judgments of variance were made no more accurately for abutting relative to separated textures. Interestingly, however, responses were overall faster for textures differing in variance when they were separated compared to when they were abutting. This is perhaps due to the clear separation boundary, which serves to delineate the regions on which to perform some regional estimation of orientation variance. A second experiment highlights the phase-insensitivity of texture segmentation, in that locating a texture edge (defined by a difference in mean orientation) in high frequency orientation-reversing stimuli can be performed at much higher frequencies than the discrimination of the same regions but with the texture contour masked. Textures that differed in variance did not exhibit this effect. A final experiment demonstrates that the phase-insensitive perception of texture borders improves with eccentric viewing relative to the fovea, whereas perception of the texture regions does not. Together, these experiments show dissociations between edge- and region-based texture analysis mechanisms and suggest a fast, sign-invariant contour extraction system mediating texture segmentation, which may be closely linked to the magnocellular subdivision of visual processing.

2.2. Introduction

Natural scenes contain contours, which separate figures from their ground or represent discontinuities in object surfaces. These are typically associated with spatial gradients in chromaticity or luminance, although contours can also be perceived as a result of spatial changes in texture in the absence of such cues. This occurs very noticeably when two abutting textures differ in the mean orientation of their textural elements (see Bergen, 1991; Nothdurft, 1991; Thielscher & Neumann, 2003). Although the strength of the border's perception is dependent on the orientation noise within the stimuli (Motoyoshi & Nishida, 2001), generally the greater the orientation contrast, the more compelling the border (Nothdurft, 1991).

This process of contour localisation allows effortless texture segmentation to occur. The rapidity of this process is highlighted in a number of psychophysical and neurophysiological findings. Motoyoshi (1999) showed that when a uniform texture was presented briefly and followed by a smaller mask, the textured area within the mask would be suppressed and the overall perception would be that of a uniform texture with a 'hole' in the centre. Importantly, however, if the texture contained an orientation discontinuity within the area of the mask, this region survived the mask's effect. Additionally, Lamme, Rodriguez-Rodriguez and Spekreijse (1999) found single-cell evidence in the awake macaque of a response enhancement at the location of an orientation contrast that occurred prior to any location within the regions of the textures themselves. The largest response latencies were recorded for regions farthest away from the contours. Similar findings were recorded by Romani, Caputo, Callieco, Schintone and Cosi (1999) using visual evoked potentials, which, together, corroborate with Motoyoshi (1999) in suggesting the workings of a texture analysis mechanism that prioritises locations of contrast relative to regions of uniform texture.

This process of edge-based texture perception is thought to be achieved first through the application of first order orientation-specific filters in the visual system, whose outputs are then transformed through a second stage non-linearity. A final stage of linear filtering is then conducted on this output with larger receptive fields than those of the first, and is sensitive to the location of any spatial gradient in the orientation-defined texture (see Bergen, 1991, for a detailed review of such a model). The output of the filtering is such that a peak response is created at the location of contrast. In other words, the process of segmentation based on orientation contrast, and potentially texture segmentation more generally, is *edge-based* (Wolfson & Landy, 1998).

The term *segmentation* is often reserved for the description of the processes of edge-based mechanisms, whereas the term *discrimination* is most often applied to instances in which an observer is able to discern two textures despite the absence of an informative boundary (Landy & Graham, 2004). These instances arise either when spatial or temporal separation of the textures prevents their segmentation, or by the presence of a texture difference that is not well suited for the filter-rectify-filter process. The mechanisms underlying this process are described as being *region-based* and require the integration of spatially distributed local signals to attain an informative statistical measurement of the composition (Dakin & Watt, 1997; Wolfson & Landy, 1998). This statistical analysis is essential to our consistent perception of visual texture more generally, as comparison on a pixel-by-pixel basis would be too unreliable and heavily dependent on various viewing conditions (e.g. viewing angle, lighting, and distance; Adelson, 2001) if we are to recognise one instance of texture as belonging to a particular category (e.g. sand, granite or woodchip). Indeed, Portilla and Simoncelli (2000) have shown that new instances of a particular texture can be successfully generated on the basis of statistical representations garnered from a sample image (e.g. from the responses of orientation and spatial frequency filters and correlations between such filters). One example of a simple image statistic that has been studied in relation to visual texture is that of orientation variance. Spatial discontinuities in orientation variance are detected poorly by a filter-rectify-filter process (Wolfson & Landy, 1998), as the average orientation on either side of the border is constant. Dakin and colleagues (Dakin & Watt, 1997; Dakin, 2001) have shown that observers' judgments of orientation variance were both accurate and flexible, being dependent on the characteristics of what is being perceived. Additionally, Morgan, Chubb and Solomon (2008) have outlined a 'dipper' function in the representation of orientation variance in visual texture that may be evidence of intrinsic noise resulting from a dedicated mechanism in the visual system for the computation of visual texture. Thus, two very distinct mechanisms appear to exist; one which prioritises the extraction of texture-defined contours, and one which estimates statistical properties within a region.

Wolfson and Landy (1998) took the approach of directly comparing and contrasting these mechanisms. They showed that when two patches of texture differed in mean orientation, then observers were more sensitive to the difference if the patches were abutting than if they were separated. When the patches differed in variance, however, no such effect was observed. This is indicative of a 2nd-order image statistic (i.e. orientation variance) being associated with a region-based mechanism; in other words not related to texture segmentation, but more related to the discrimination of a texture's *appearance.* A trend was also found in a subset of their participants that discriminations based on differences in variance were actually greater when the patches were *separated.* This was theorised to be due to the role that a separating boundary might have played in delineating the areas on which to perform the regional analysis (in this instance, variance estimation).

The aim of the first experiment was to replicate the findings of Wolfson and Landy (1998) but also to correct for a potential confound in their experiment: their stimuli were composed of randomly positioned line segments and thus were allowed to overlap. This created a larger quantity of "line crossings" in textures of high variance, a cue that observers could have used to discern variance independently of any computation of orientation statistics (Julesz, 1981). Although Wolfson and Landy (1998) discussed this and dismissed it, no experimental evidence has yet to be presented which directly addresses this point, and data collected from experiments using such stimuli may reveal more clearly the workings of region-based texture analysis mechanisms. This confound is prevented in the present study by using structurally placed Gabor patches that do not overlap. It is predicted that for odd-one-out judgments based on a difference-in-mean, accuracy would be higher for abutting patches than for separated patches of texture. Conversely, no effect is predicted for discrimination based on a difference-in-variance. In addition, the first experiment extends Wolfson and Landy's (1998) findings by measuring reaction times (RTs) on the same task, predicting that when accuracy is greatest for abutting textures relative to separated textures in the case of detecting a difference-in-mean, RTs would also be shorter. Conversely, no decrease in RTs for detecting a difference-in-variance is expected, although RTs are expected to be faster for separated textures, as the separation may serve to clearly define the regions on which to perform regional estimations of variance.

2.3. *Experiment 1 – oddity detection task for differences in orientation mean and variance*

2.3.1. Methods

Participants

Ethical approval for this experiment, and all subsequent experiments reported in this thesis, was obtained from the Psychology Research Ethics Committee at Durham University. 20 participants (11 male, 9 female) took part in all conditions of this study. All subjects had normal or corrected-to-normal vision. All gave written informed consent to take part and were compensated financially for their time.

Stimuli

The display monitor was viewed at a distance of 41cm (subjects rested their head on a chin rest). Stimuli were presented on the uniform grey background of a gamma-corrected ViewSonic 17'' (1254 x 877 pixels) colour monitor positioned on its side and driven by a Cambridge Research Systems VSG 2/5 Graphics System.

The experimental stimuli consisted of three regions of texture presented in a vertical alignment 0.5° to the right of fixation (a fixation cross subtending 0.3° x 0.3°), as measured from the edge of the stimuli. Each was constructed of 240 evenly positioned Gabor patches in a 24 (across) x 10 (down) rectangular arrangement. Each individual Gabor patch had a spatial frequency of 3 cycles/degree and was generated with a common cosine phase such that the pivotal centre of each Gabor patch contained a dark band. Each measured 0.5° in diameter and was separated from its neighbours by a distance of 0.3°. Thus, each patch of texture covered a total area of 27.0° x 17.7°. Gabors had a 100% luminance contrast. It was decided that the locations of the Gabors were to be fixed rather than randomly distributed to avoid any potential confounds of "line crossings" that would occur with an increase in variance.

The orientation of each individual Gabor patch was independently drawn from a Gaussian distribution of a particular mean (μ) and variance (σ^2) associated with its relative texture patch. In all trials, a 'pedestal' patch was created with the parameters $\mu=\alpha$ (where α is a randomly determined orientation between 0° and 360° cycle) and $\sigma^2= 5^2$. In a difference-in-mean trial, a second patch was created with the parameters $\mu = \alpha \pm x$ and $\sigma^2 =$

5^2 (where *x* is the degree of orientation difference for that trial, i.e. x = 2, 5, 8, 11, 15 or 20). In a difference-in-variance trial, the second patch was created with the parameters $\mu=\alpha$ and $\sigma^2 = (5 + y)^2$ (where *y* determines the degree of variance difference for that trial, i.e. y = 0.75, 1.25, 2, 3.75, 5 or 7.5, thus creating variance differences of 8.1, 14.1, 24, 51.6, 75 and 131.3). In both types of trial, the third patch was constructed with the same parameters as either the pedestal patch or the second patch such that the odd texture was equally likely to be the pedestal patch or not.

Procedure

Trials were completed in two blocks, each defined by the adjacency of the textures: in one block (the *abutting* condition), the patches were adjoining such that the distance between each patch of texture equated the distance between the Gabor patches within each texture patch (0.3°). In another block (the *separated* condition), the patches of texture were separated vertically by a distance of 0.5° between each one, thus creating the impression that there were 3 isolated patches of texture one atop the other (see **figure 2.1**). Additionally, on each trial, the position of the stimuli was randomised such that the whole display was equally likely to be shifted 0.9° either above or below fixation level or not at all. This was designed to prevent the observer from anticipating the exact location of the textures and thus from pre-empting the position of the border.

In each trial, two of the textures were given the same appearance (by drawing their composite orientations from the same Gaussian distribution as one another). The third texture (the 'odd' texture) was equally likely to appear at the top or bottom of the arrangement, but never in the centre. This patch of texture was made conspicuous compared to the others by changing either the *mean* or *variance* of the distribution from which its orientations were drawn. See **figure 2.1** for examples of each type. For each type of texture difference (difference-in-mean/ variance), there were 6 difference levels. They were, for a difference in mean, 2, 5, 8, 11, 15 and 20 degrees, and, for a difference in variance 8.1, 14.1, 24, 51.6, 75 and 131.3. For each block of trials corresponding to each condition of texture adjacency (abutting/separated) there were 2 conditions of texture difference (difference-in- mean/difference-in-variance) x 6 difference levels x 8 repetitions. Thus there were 96 trials per block per subject. It should be noted that the order of trials was randomized within each block, ensuring that for each given trial subjects could not anticipate the type of parameter that was to define the “odd” texture from the rest (i.e. whether it was a difference-in-mean or difference-in-variance). This adds a particular level of task uncertainty for the subjects and so reduces any variation in how they might approach the different discrimination types; in other words, they are not able to adopt a preferred strategy over the course of one block of trials that would selectively appeal to discriminating difference-in-mean or difference-in-variance textures. The presentation order of the blocks was also counterbalanced across participants.

Thus, there were four conditions of the experiment, with each containing 6 levels of difference

1. patches are abutting, with a difference in mean
2. patches are abutting, with a difference in variance
3. patches are separated, with a difference in mean
4. patches are separated, with a difference in variance

Figure 2.1 shows examples of each of these four types of stimuli. The stimuli were presented for 5s or until a response was made. A fixation cross appeared 2s prior to the onset of the stimuli and remained until a response was made. Subjects were instructed to fixate the cross whenever it was present and to respond as accurately and as quickly as possible in a two-alternative forced choice paradigm by pressing the appropriate key (top or bottom) to indicate the odd texture. Responses were made on a 5-button Cedrus Response Box (Cambridge Research Systems) with the subject’s right hand. The next trial would not begin until a response had been collected and a 2s inter-trial interval was included. Accuracy and reaction times were measured. Stimuli were presented to the right of fixation.

2.3.2. Results

Figure 2.2 shows the data from experiment 1. To reduce the noise in the data, the lowest two difference levels were excluded for both the difference-in-mean and difference-in-variance conditions, as participants’ accuracy did not rise above chance performance, leaving 4 levels of difference in each condition. RTs that exceeded 4 seconds were removed from analysis along with any remaining values that fell outside 2 standard deviations of the mean per condition per subject. This removed no more than 9.5% of all cases. Planned comparisons were performed by isolating either the mean or variance discriminations for either the accuracy or RT measurements. Thus, four 2 x 4 within-subject ANOVAs were carried out on these data with the factors of texture adjacency ('abutting' vs. 'separated') and level of difference, respectively. Each of these will be discussed in turn. In all of these analyses, the main effect of difference level reached significance (all F values>20) and thus is omitted in all cases from the report to aid concision.

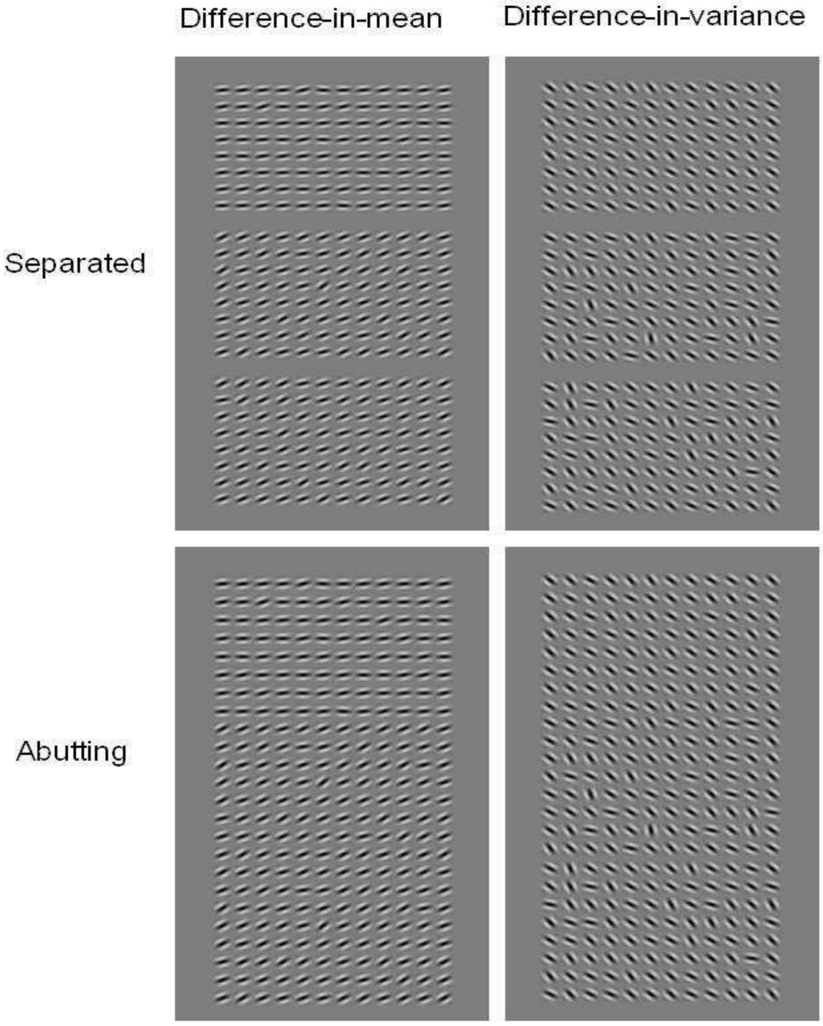

Figure 2.1. Replications of the stimuli used in the four conditions of experiment 1. Each window shows 3 equally sized textures either *abutting* or *separated* vertically from one another. The task required observers to indicate the "odd-one-out". The odd texture would either be due to a *difference-in-mean* or a *difference-in-variance* in the orientation statistics used to generate each texture. Thus, in the figure, the top texture in each of the quadrants is the odd-one-out. Stimuli were presented slightly to the right of a small fixation cross. Note that the textures in the illustration are a reduction in size of the actual stimuli used.

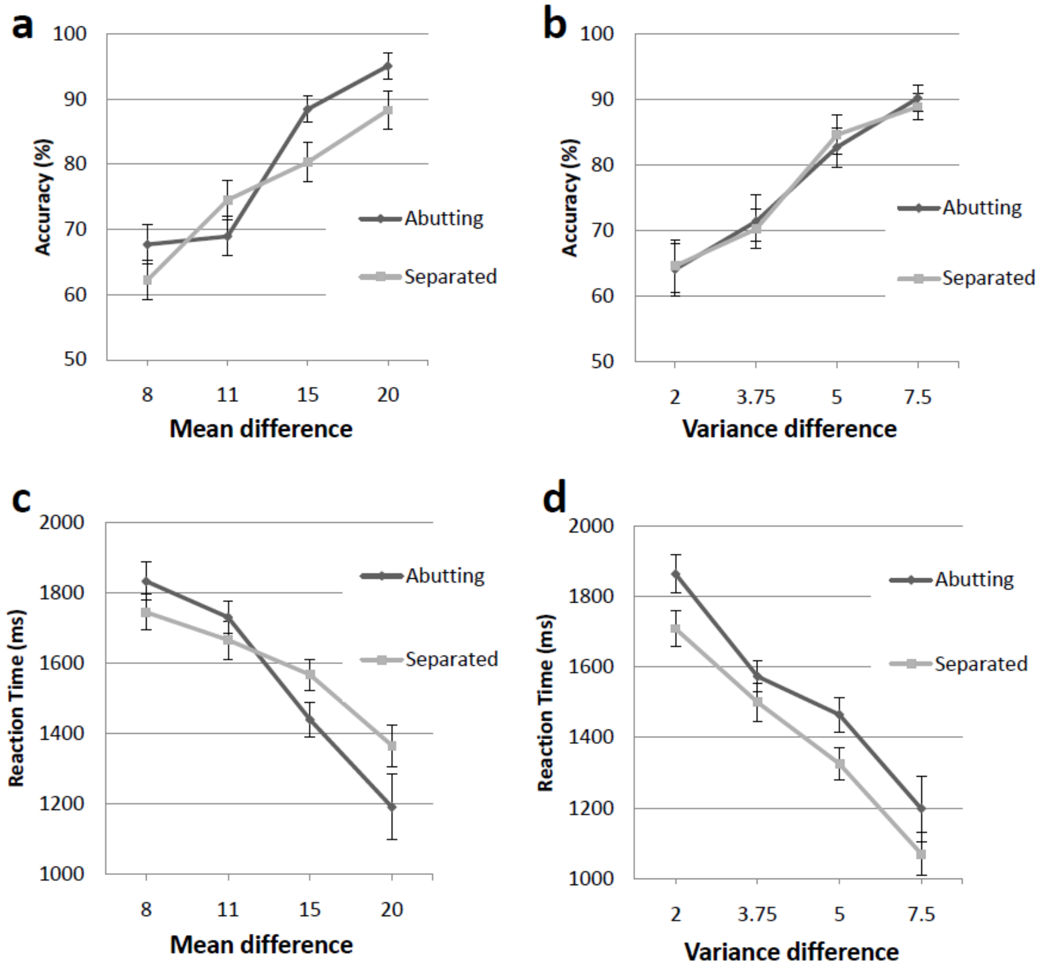

Figure 2.2. Results from experiment 1. **a & c:** accuracy and RT data for difference-in-mean discriminations for either abutting or separate textures. **b & d:** accuracy and RT data for difference-in-variance discriminations for either abutting or separated textures. All are shown as a function of increasing difference level. Error bars show +/- 1 SEM with between-subject variance omitted.

Accuracy data for difference-in-mean stimuli (**figure 2.2a**): No significant effect of texture adjacency emerged ($F_{(1,19)}$ = 3.365, p=0.082), but there was a significant interaction between the two main effects ($F_{(3,57)}$ = 3.159, p = 0.031). As subsequent paired t-tests showed, no difference in accuracy was found between odd-one-out judgments of abutting and separated textures at the lower levels of difference, specifically 8-degrees ($t_{(19)}$ = 1.20, p=0.245) and 11-degrees ($t_{(19)}$ = 1.555, p = 0.137). Higher accuracy rates, however, were found for judgments of abutting- relative to separated-textures at 15-degrees ($t_{(1,19)}$ = 2.106, p = 0.049) and 20-degrees ($t_{(1,19)}$ = 2.963, p = 0.008) with differences of 8.2% and 6.8%, respectively.

Accuracy data for difference-in-variance stimuli (**figure 2.2b**): No significant effect of texture adjacency emerged ($F_{(1,19)}$ = 0.142, p=0.710) and there was no evidence of an interaction ($F_{(3,57)}$ = 0.112, p = 0.947). Note the extreme similarity between the plots in **figure 2.2b** compared to those in **figure 2.2a**. These data indicate that subjects' accuracy for odd-one-out judgments based on orientation variance did not improve with abutting textures relative to separated textures.

RT data for difference-in-mean stimuli (**figure 2.2c**): No significant effect of texture adjacency ($F_{(1,19)}$ = 0.213, p=0.649) was found, but a significant interaction emerged ($F_{(3,57)}$ = 3.334, p = 0.026). Subsequent paired sample t-tests revealed no difference in RT for odd-one-out judgments of abutting and separated textures at any level of the difference in mean orientation (all p values>0.1), however, although **figure 2.2c** clearly shows a trend that is consistent with the accuracy data (**figure 2.2a**); at higher difference levels (15- and 20-degrees) responses were faster when the textures abutted compared to when they were separated, but this effect was not present at lower levels of difference, which in fact show the opposite effect. It is important to note at this point that there is no evidence that subjects showed a speed-accuracy trade-off; neither did they make more accurate judgments by delaying their RT and nor did they make quicker decisions through sacrificing accuracy.

RT data for difference-in-variance stimuli (**figure 2.2d**): A significant effect of texture adjacency ($F_{(1,19)}$ = 4.437, p=0.049) was found, in which subjects were, on average, *quicker* by 124.01 ms to judge *separated* textures than abutting textures. No significant interaction emerged ($F_{(2.061,39.16)}$ = 0.190, p = 0.83; Greenhouse-Geisser corrected). Again, no evidence of a speed-accuracy trade-off was found.

2.4. Interim discussion

The results support Wolfson and Landy's (1998) conclusions by showing that accuracy increased for judgments made based on mean orientation when the textures abutted compared to when they were separated (although only at the higher levels of difference). This was further supported by the finding that when accuracy was greater for abutting textures, RT was also faster, rather than subjects trading speed for accuracy. Contrarily, when subjects weren't significantly more accurate, they were, if anything, slower. This is consistent with the detection of a difference in mean orientation being well suited for a fast contour-detection system that has been outlined in both the psychophysical and neurophysiological literature (e.g. Motoyoshi, 1999; Lamme et al., 1999; Romani et al., 1999). Importantly, however, there was no such elevation in accuracy for judgments based on variance; performance was remarkably similar between the two conditions. This is because on either side of the boundary between the textures, the average orientation is the same and so a filter-rectify-filter process tuned to orientation would elicit very weak responses at this location of contrast. Perhaps the most intriguing finding, however, is that judgments based on differences-in-variance were made more rapidly for separated textures compared to those abutting. Wolfson and Landy (1998) briefly discussed that it might be expected that subjects would perform better on judgments of variance if the patches were separated, suggesting that the separation of the textures serves to clearly delineate the regions to be discerned. The flexible "region of integration" of orientation signals associated with variance estimation (Dakin & Watt, 1997; Dakin, 2001) may indeed benefit from the clear border definitions offered by the separation of the textures. In the present experiment, however, this is an effect which manifested clearly in measurements of RT but not accuracy. Given different instructions, and perhaps limited stimulus presentation durations, subjects may have shown significantly greater accuracies in the separated relative to the abutting condition.

The notion of a fast edge-based mechanism of texture segmentation is explored further in experiments two and three. Specifically, is rapid texture segmentation by orientation contrast a result of phase-insensitive magnocellular processes? The cells of the magnocellular subdivision of visual processing (M cells) exhibit high temporal and low spatial sensitivity, as well as being adept at discriminating motion but being blind to colour identity. This is in stark contrast to the slower parvocellular system, which is more attuned to spatial detail and chromatic information (Schiller & Logothetis, 1990). One important aspect of M cells is that they have the capacity to detect chromatic contrast despite their colour-insensitivity (Saito, Tanaka, Isono, Yasuda & Mikami, 1989). So an intriguing question is whether the border that is perceived between two regions of texture that differ in mean orientation is generated by a magnocellular mechanism that signals orientation contrast without signalling the identity of the orientations themselves.

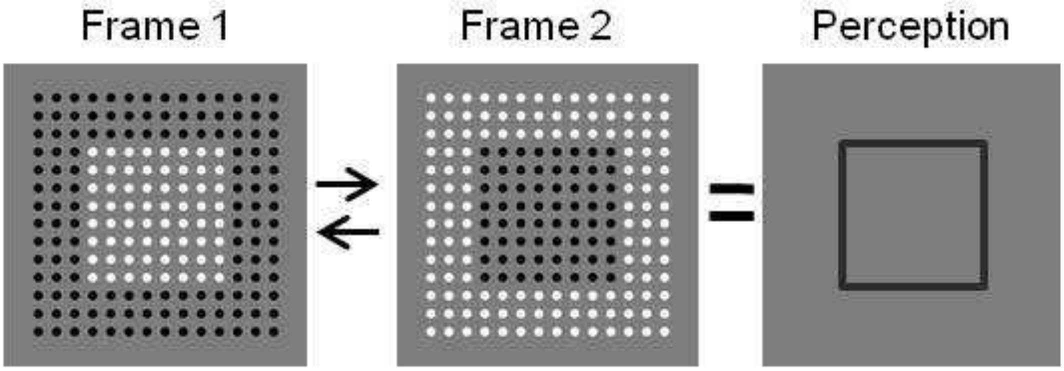

Figure 2.3. Example of flicker-defined form. When the two frames shown above are presented in high frequency alternation, an implicit border can be seen between the inner and outer regions despite it being impossible to tell at any given time which is dark and which is light. Thus, when an intervening black border is placed between the two regions, the viewer cannot perceive a difference between the two regions at such high frequencies. Note that the perception shown in the figure is just an illustration, black lines are not seen.

The second experiment investigated this possibility using “flicker-defined form”. Using such stimuli, Rogers-Ramachandran & Ramachandran (1998) demonstrated that when two abutting fields of dots (one consisting entirely of white dots, the other black; see **figure 2.3**) are flickered in counter-phase above a particular frequency (roughly 7 Hz), the resultant perception is that of an implicit border separating two indistinguishable regions. Up to roughly 15 Hz (depending on various stimulus attributes), this perception of the border remains despite no conscious access to the surface information (i.e. which side is black, which is white). This particular type of stimulus is thought to isolate the phase-insensitivity of magnocellular mechanisms and has been used as a diagnostic test for its improper functioning (e.g. in cases of dyslexia, Sperling, Lu, Manis & Seidenberg, 2003). The absolute effectiveness of this type of stimuli in isolating magnocellular functioning, however, is slightly contentious (see Discussion [section *7*], and Skottun & Skoyles, 2006).

Assuming that texture segmentation is governed by the magnocellular system, if the orientations of two textures are reversed above a particular frequency, the texture border should be visible at a higher frequency than is possible to discern the regional texture information (i.e. the identity of the orientation either side of the border). This was assessed in the second experiment by requiring observers to locate a texture border defined by a difference in mean orientation in both the presence and absence of a coincidental black border, which masks the perception of an implicit contour that may be formed between the two regions. Thus, it is assumed that two processing mechanisms exist: *edge*- and *region*-

based texture analysis, and that these processes can be dissociated in the case of the difference-in-mean condition by showing that edge-based mechanisms operate at a higher speed than their region-based counterpart. As a control, no such dissociation should exist for textures differing in variance, as judgments based on this statistic are not governed by the fast contour-extraction (edge-based) mechanism.

2.5. Experiment 2 – flicker-defined-form for differences in orientation mean and variance

2.5.1. Methods

Participants

3 naive observers (2 male) as well as the author (male) took part in all conditions of the experiment. All had normal or corrected-to-normal vision and gave their full written informed consent.

Stimuli

This display set-up and equipment used were identical to the previous experiment. A black fixation cross was present in the centre of the screen before and after the presentation of the stimuli. The experimental stimuli consisted of a lattice of 20 x 20 uniformly positioned Gabor patches (each measuring 0.6° in diameter and separated from its neighbours by 0.2°). In total, the lattice measured 14.8° in width and in height. All Gabor patches had a spatial frequency of 3 cycles/degree and were each given a randomly determined phase from the full 0-360° cycle. This phase randomisation ensured that integrated contours formed by neighbouring Gabors along a common path were not detectable by a 1st-order mechanism. Gabors had a 100% luminance contrast.

The textures' statistics were determined by independently drawing each composite orientation value from a Gaussian distribution with a particular μ and σ^2. For one half (the "pedestal" texture), orientations were drawn from a distribution with a randomly determined μ for that trial and a σ^2 of 10^2. In the difference-in-mean sessions, the remaining half of the lattice (the second texture) would differ in μ with a magnitude of 90°. In the difference-in-variance sessions, μ would be the same but the second texture would differ in σ^2 by a magnitude of $+26^2$.

Procedure

Trials were divided into four blocks. In each block, the subject's task was to identify how the lattice was bisected (vertically vs. horizontally), which was determined randomly with equal probability on each trial. Subjects completed two blocks of trials in which the lattice was divided by a difference in the mean orientation between the two textures (see top row of **figure 2.4**) and two in which the lattice was divided by a difference in the orientation variance (see bottom row of **figure 2.4**). For each of these types of stimuli, one of the two blocks of trials included the addition of two intervening black borders (separators) dividing the lattice both vertically and horizontally at the locations at which the change in orientation statistics would coincide (see right column of **figure 2.4**). The separators measured only 1 pixel in thickness and 15.1° in length. This masked the effect of any subjective contour that may have been brought about by an orientation contrast. Unlike the previous experiment, the stimuli were to be placed in the centre of the visual field, so the decision was taken in this experiment to use black borders to separate the textures, as opposed to the spatial separation used in the first experiment, to control for effects of eccentricity. Thus, the 4 trial blocks were, with the order being counterbalanced across participants:

1. Difference-in-mean, abutting
2. Difference-in-mean, separated
3. Difference-in-variance, abutting
4. Difference-in-variance, separated

Figure 2.4 shows illustrations of the stimuli used in each of the above conditions. In each trial, two frames of stimuli were generated and presented in alternation at different levels of frequency. The first frame was generated by constructing the textures through the methods described above. The second frame, however, was generated by exchanging the orientation statistics of the two halves of texture. Note, however, that the two frames were not simply mirror images of one another; the two frames contained exactly the same orientation values, though their positions were randomised within each texture. This randomisation was designed to prevent observers from using the Gestalt principal of "common fate" to infer how the lattice was divided. For example, by witnessing a particular cluster of orientations moving together in a particular direction (e.g. left to right, or top to bottom), the way in which the textures were bisected (horizontal vs. vertical) could be inferred. This ensured that the task could only be performed on the basis of statistical computations of the orientations.

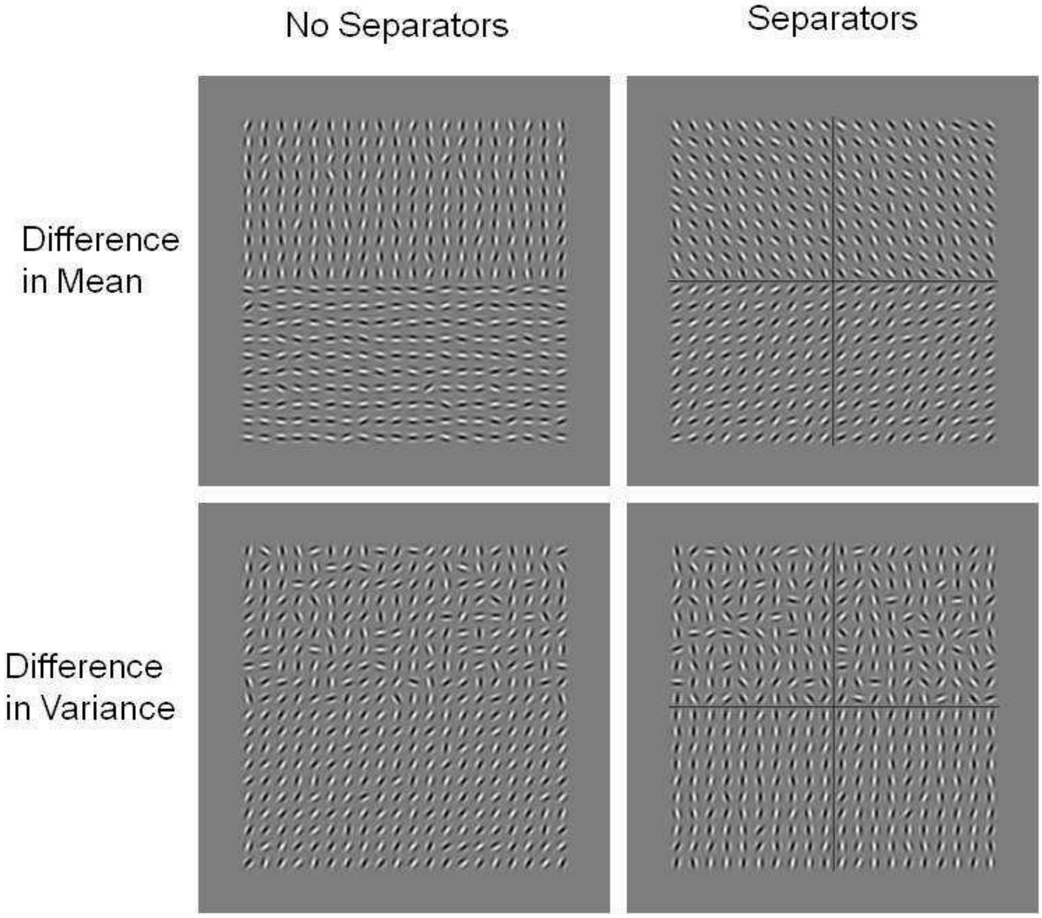

Figure 2.4. Example stimuli used in experiment 2 for each of the four conditions. In each case, the texture is divided horizontally.

Figure 2.5 shows an illustration of the sequence of events for one trial. The subject fixated centrally on the black cross before each trial. Each trial was preceded by a warning tone, after which (250ms) the fixation cross would disappear and a mask would appear at the centre of the screen consisting of a lattice of equal size as the stimuli but consisting of randomly determined orientations. This mask would last for 30ms before being replaced by the stimuli. The stimuli would last on screen for 800ms, with the two textures continually flickering (appearing to change positions) at a particular frequency until the offset of the stimuli. This was followed by a second mask identical to the first one.

Following this (500ms), a second tone signalled to the participant to make a response. Participants made a two alternative forced-choice decision, pressing one key if they thought the lattice was divided vertically and pressing another if they thought the lattice was divided horizontally. Subjects were encouraged to guess if they were unsure. Responses were made on a 5-button Cedrus Response Box (Cambridge Research Systems). No time limit was set for the collection of the response, but the next trial would not begin until a response had been collected. A 1s interval was included between the period of collecting the response and the warning tone for the following trial, during which only the fixation cross was present.

Each block of trials consisted of 4 interleaved one-down two-up staircases that each began at a low frequency (4 Hz) that increased in frequency following 2 correct responses and decreased following 1 incorrect response. This method estimates the 70.7% correct level on the subject's psychometric function. The magnitude of the frequency increment/decrement was fixed at 1 Hz for the first 10 trials in each staircase, and then at 0.4 Hz in all subsequent trials. The staircases each terminated after 14 reversals in performance.

2.5.2. Results

Figure 2.6 shows the data collected from experiment 2. Data were collected at frequency levels (measured in Hz) at which subjects' accuracy at successfully judging the division of the lattice rested at 70.7%. Each staircase estimated this threshold by averaging the final 6 reversal points in performance (after a total number of 14 reversals). The values shown in **figure 2.6** are the average threshold estimates taken from the four staircase procedures in each condition for each subject.

It is clear from **figure 2.6** that, for the difference-in-mean condition, all subjects could identify the division of the abutting textures at a higher frequency compared to those that were separated with black lines. Importantly, for the difference-in-variance condition, no similar effect was observed; subjects either performed comparably in the abutting and separated conditions or reached a higher frequency, albeit marginally, in the separated condition. Because these difference-in-variance results were not found consistently for all 4 subjects, however, no strong inferences can be made as to whether they truly suggest better performance in the separated relative to the abutting condition. Experiment 1, however, does suggest that this is possible.

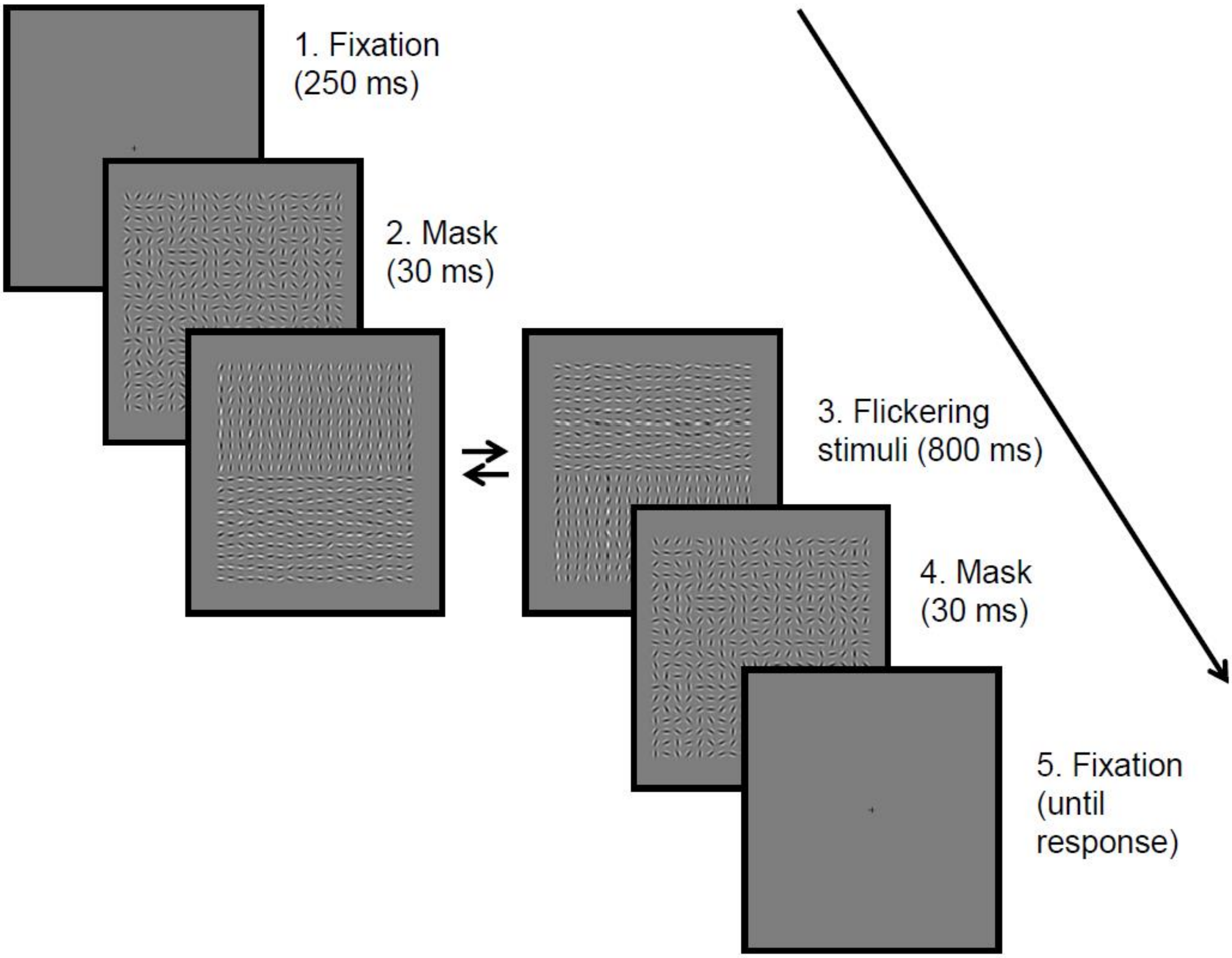

Figure 2.5. The presentation of the stimuli in each trial for experiment 2. The illustration shows an example of a trial in which a lattice was presented with a difference in mean orientation between the left and right regions. For the duration that the stimuli were displayed (in stage 3 of the illustration), the two frames were presented in alternation at a particular frequency, as indicated by the double arrow. Subjects were required to indicate whether the lattice was divided horizontally (as in this case) or vertically.

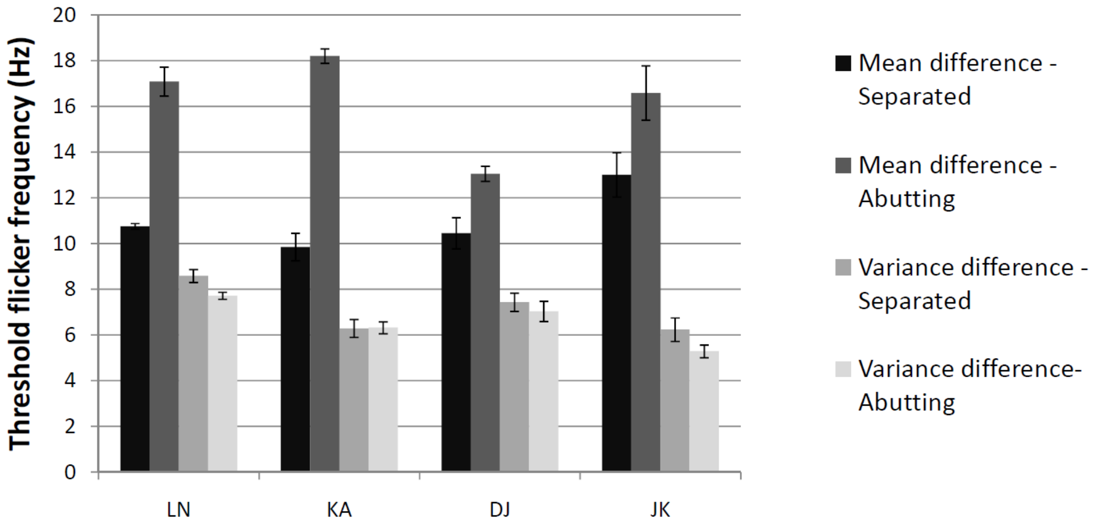

Figure 2.6. Results from experiment 2. Estimates of threshold flicker frequency for each condition per subject. Values show the frequency level at which subjects performed at 70.7% in judging the difference between two textures either based on a difference-in-mean or a difference-in-variance and either with or without the inclusion of coincident black borders at the locations of contrast. Error bars represent +/- 1 SEM, taken from 4 staircase estimates of threshold.

Most importantly, however, the present results strongly indicate that, like the two textures composed of black and white spots used by Rogers-Ramachandran & Ramachandran (1998), when two textures composed of orientation signals have a difference in mean orientation, the perception of an implicit border between these regions can be dissociated from that of the actual orientations on either side of the border. The fact that the perception of the border persisted to a higher frequency than that of the regional qualities is potential evidence of magnocellular processing underlying the perception of orientation contrast.

2.6. Interim discussion

Evidence that texture segmentation may be the result of magnocellular processing has also come from findings that the rapid detection of a figure defined by texture contrast increases with non-foveal viewing (Joffe & Scialfa, 1995; Saarinen, Rovamo & Virsu, 1987). Generally, most visual discriminations diminish with peripheral viewing, but one characteristic of peripheral vision relative to central vision is the abundance of M cells in comparison to P cells in the retina (de Monasterio & Gouras, 1975).

The third experiment aims to show that the perception of the border alone defined by a difference in mean orientation is more readily perceived in the periphery compared to the fovea, whereas this is not true for the perception of the regions of texture that constitute the border.

2.7. Experiment 3 – flicker-defined-form for differences in mean orientation in the fovea and periphery

2.7.1. Methods

Participants

The participants were those from experiment 2.

Stimuli

The stimuli and set-up were the same as the previous experiment, although only the "difference-in-mean" conditions (with and without the separators) were used. Also, the lattice contained 10 x 10 uniformly positioned Gabor patches, rather than 20 x 20, thus measuring 8.0° in width and in height. For blocks of trials in which the subjects viewed the stimuli peripherally, the lattice would appear 5.6° to the left of the fixation cross (as measured from centre-to-centre).

Procedure

The general procedure was the same as the previous experiment, except that subjects completed 2 blocks viewing the stimuli centrally and 2 blocks viewing the stimuli in the periphery. As in the previous experiment, one of the blocks in each of these conditions would present the stimuli with the black separators (8.0° in length) that bisected the texture horizontally and vertically. Thus, the 4 trial blocks were, with the order being counterbalanced across participants:

1. Foveal viewing, abutting
2. Peripheral viewing, abutting
3. Foveal viewing, separated
4. Peripheral viewing, separated

2.7.2. Results

Thresholds were estimated in the same way as in the previous experiment. **Figure 2.7** shows these data. For all subjects, in the abutting condition frequency thresholds were higher when the stimuli were viewed peripherally compared to when they were viewed foveally. Contrastingly, in the separated condition, for two of the subjects (KA and JK) thresholds were comparable and for the other two (LN and DJ) performance was poorer (lower frequency threshold) in the periphery. On no occasion did performance improve in the periphery relative to the fovea for separated textures, whereas it did so consistently for abutting textures.

As there are fewer M cells found in the fovea compared to the surrounding retinal area (de Monasterio & Gouras, 1975), this finding is consistent with the notion that the ability of the visual system to signal a texture border defined by orientation contrast is mediated by the magnocellular system, whereas the perception of the regions of the texture is governed by a separate, possibly parvocellular, system.

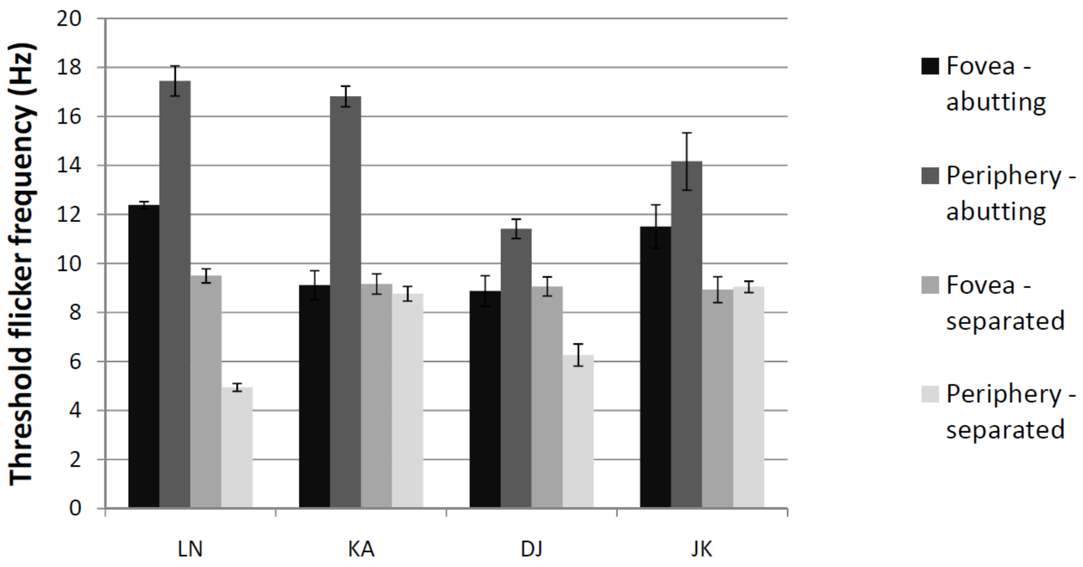

Figure 2.7. Results from experiment 3. Estimates of threshold flicker frequency for each condition per subject. Values show the frequency level at which subjects performed at 70.7% in judging the difference between two textures that differed in mean orientation either in the fovea or at 5.6° in the periphery and either with or without the inclusion of coincident black borders at the locations of contrast, taken from 4 staircase estimates of threshold.

2.8. General discussion of study one

The results gained from these three experiments highlight a fast, edge-based texture segmentation mechanism responsible for signalling the location of a border between two textures defined by a difference in mean orientation. Performance was more accurate and more rapid for abutting textures relative to separated ones (at high levels of difference). Additionally, when the textures were rapidly phase-reversed, the implicit border formed by the orientation contrast could be readily perceived at a much higher frequency than the perception of the orientations either side of the border. Furthermore, experiment three showed that the perception of such contours improved with non-foveal viewing, whereas the perception of the regions constituting such contours did not. This corroborates previous work that has shown texture segmentation more generally to improve in the periphery (Joffe & Scialfa, 1995; Saarinen *et al*, 1987), but additionally demonstrates that it is the perception of the texture *contour* alone that underlies this effect. These dissociations suggest separable mechanisms of edge- and region-based texture perception, and potential magnocellular involvement in the former, as some M cells have been shown to signal properties such as chromatic contrast and luminance contrast whilst remaining insensitive to the identity of the stimulus properties either side of the border (Saito *et al,* 1989; Rogers-Ramachandran & Ramachandran, 1998), and their relative number increases in the periphery.

It is partly questionable, however, as to whether flicker-defined-form does indeed isolate magnocellular processing. Indeed, this can only be speculated, as no study has yet recorded selective activity from an M cell in response to such stimuli. This principle is merely the result of observing the similarities between the perception of flicker-defined-form and the responses of M cells (i.e. they improve with lower spatial frequencies, are phase insensitive and improve or are more numerous in the periphery, Rogers-Ramachandran & Ramachandran, 1998; Schiller & Logothetis, 1990; Saito *et al,* 1989; de Monasterio & Gouras, 1975). Despite this, Skottun and Skoyles (2006) have speculated that the disparity between flicker frequencies associated with the perceptions of surface and contour information in flicker-defined-form is too large to imply a distinction between parvo- and magnocellular processes. Rather, it is perhaps indicative of a cortical distinction, rather than a subcortical one. Lamme and colleagues (1999) and Romani and colleagues

(1999) certainly found a temporal enhancement of the processing of texture edges relative to homogenous regions in the *cortex.* It is not clear, however, whether this represents a cortical distinction independent of magno- and parvocellular processes or simply reflects the relative activity of the pathways which project onto the cortical regions in question. Clearly more physiological recording needs to be carried out if we are to answer this question fully and understand the relationship between the perception of flicker-defined-form and the magnocellular pathway. Whether or not the method of using flicker-defined-form in experiments 2 and 3 indeed isolated a magnocellular process of texture segmentation is clearly important, but regardless of the answer to this question, it does not detract away from the observation that texture segmentation can be achieved through rapid, phase-insensitive processes, regardless of their neural substrates.

On a related note, in addition to orientation contrast, texture segmentation based on spatial frequency would also be an interesting consideration for phase-insensitive texture segmentation. Bergen & Adelson (1988) have shown how the automatic segregation of textures based on element shape (‘X’s vs ‘L’s) diminishes following the equating of spatial frequency content across the border. Such perceptual computations may be undertaken by early mechanisms in the visual brain, found in the lateral geniculate nucleus and layer 4 of the primary visual cortex (Landy & Graham, 2004), and so may be susceptible to the effects shown in the present study.

In comparison to the rapid edge-based texture segmentation mechanism, some texture discriminations require a more regional, statistical analysis and are not well suited for the processing of edge-based mechanisms. This was shown in the case of orientation variance judgments in the present study. In experiment one, unlike judgments of mean orientation, those of orientation variance were not more accurate when the patches were abutting compared to when they were separated. This is in line with Wolfson and Landy’s (1998) findings and their interpretation of edge- and region-based texture analysis mechanisms, but the present study removed the confound of an increase in the quantity of line crossings in textures with greater variance. Interestingly, however, in the present experiment, judgments were actually more rapid when the patches were separated than when they were abutting. Wolfson and Landy (1998) offered the possibility that the separation serves to clearly delineate the regions on which to perform the regional analysis. This is certainly a possibility, given that the work of Dakin and colleagues (Dakin & Watt, 1997; Dakin, 2001) has outlined a flexible region of integration which is applied by an observer when required to perform spatial estimations of variance. In experiment one, when the textures weren't clearly separated, observers may have applied regions of integration that were inaccurately scaled to the sizes of the patches of texture. Given the relatively long presentation duration of the stimuli, however, observers were able to successfully rescale the regions, resulting in longer RTs than in the separated condition but no reduction in accuracy. Given different instructions, and perhaps limited stimulus presentation durations, subjects may have shown significantly greater accuracies in the separated relative to the abutting condition.

This finding was not entirely observed in experiment 2, however, which found that only two out of the four observers reached higher flicker frequency thresholds in the separated condition relative to the abutting condition for textures that differed in variance. The remaining two showed comparable thresholds in the two conditions. Why then does experiment 2 not clearly corroborate the findings of experiment 1? It is important to outline some differences between the experiments which may have accounted for this. One explanation is that it is only spatial *separation* that facilitates discriminations of orientation variance, rather than the presence of a coincidental black border between the textures. Alternatively, the black borders used in the second experiment not only separated the textures, but also bisected each one orthogonally, essentially dividing each texture further into two regions (see **figure 2.4**, bottom right). This may have impeded the process of orientation variance estimation, which would have worked in opposition to the facilitation effect provided by the separation of the textures, thus resulting in an overall null or very weak effect. Additionally, the stimuli in experiment 2 were fixed (i.e. appeared always in the same location) whereas those in experiment 1 were randomly positioned in one of 3 locations on each presentation. This may have increased the observer's uncertainty as to where to apply the region of integration. These are only speculations, however, and regardless of the interpretation, there is no evidence from either of these experiments of edge-based facilitation for discriminations of variance.

One peculiarity of the results which deserves to be addressed, however, is the finding that in experiment three, for foveally viewed stimuli, higher frequency thresholds for abutting relative to separated textures were found only in two of the four observers. This does not at first sight concur with the findings from the previous experiment, which found with similar stimuli higher frequency thresholds in all subjects for the abutting relative to the separated condition. One important discrepancy between these two experiments, however, was the reduction in the size (by half) of the stimuli from experiment two to experiment three. This was done to allow for more sensitive placement of the stimuli both in the centre and in the periphery of the visual field. In doing this, the stimuli in experiment three were more concentrated in the centre of the visual field for the fovea condition, and thus did not extend into the periphery as much as in experiment two. If we are to take the conclusions from the third experiment that texture segmentation improves in the periphery whereas discrimination does not, we may also speculate that the relevant results in experiment two were more evident due to the fact that the stimuli extended into the periphery, thus aiding segmentation and not discrimination.

More generally, however, these results highlight a dissociation between the processing of border information and that of surface properties, a distinction which has been at the centre of both psychophysical and neuropsychological interest for several decades. Indeed, the macro-geometry (i.e. shape) of an object appears to be processed by a separable anatomical system to that of the same object's surface properties (e.g. colour or texture; Cavina-Pratesi, Kentridge, Heywood & Milner, 2010b). Furthermore, neuropsychological cases are often presented in which a patient is unable to reliably distinguish surfaces based on some featural property but can nevertheless locate a border defined by that same property, or vice-versa. Cases of achromatopsia (cortical colour-blindness) are a striking example of this, in which the ability to use colour-opponent mechanisms is retained to the extent at which chromatic contrast can be detected, but not colour identity (Kentridge, Heywood & Cowey, 2004). So far, however, no comparable evidence has been presented for texture processing, likely to be a consequence of the highly specific brain-damage that would be required to produce such selective visual dysfunction. Nevertheless, the results gained from the present study lend well to the prediction that, at least theoretically, this could be a possibility.

3. Study Two – 2nd-order segmentation without human "V4"

3.1. Abstract

Texture *segmentation* is a rapid perceptual process, allowing object and surface boundaries to be effortlessly detected. It is currently unclear whether this is achieved in early cortical areas or whether it necessitates the region referred to as human V4. The present report presents a single case study of patient MS, whose bilateral occipitotemporal damage includes the putative human V4 area, yet whose early visual cortex is spared. As shown in these experiments, MS can accurately locate a target defined by an orientation contrast to its background, even with considerable orientation noise. Importantly, his performance was significantly reduced when the texture contours were masked by black borders (thus preventing segmentation and requiring regional discrimination), indicating that he retains a functional texture segmentation process. Additionally, when the sign of the orientation contrast was reversed at a temporal frequency of 12.5Hz, MS could nonetheless segment the target despite being unable to discriminate the target and background regions. This is an effect easily demonstrated in normal observers, and is important in isolating the rapid, phase-insensitive component of texture segmentation. Human area V4, therefore, is not necessary for texture segmentation.

3.2. Introduction

Texture segmentation, also termed 2nd-order segmentation, is the process by which the visual system rapidly detects a discontinuity in some 2nd-order property, allowing the localisation of complex boundaries between objects and surfaces. This is strongly believed to be edge-based (i.e. originating at the locations of contrast), in comparison to the region-based analyses that allow texture *discrimination.* One widely studied form of segmentation is that resulting from an orientation contrast; when two regions of orientations abut and differ in the mean orientation of their textural elements, a contour is effortlessly seen between the two regions. This process of edge-based texture perception is thought to be achieved first through the application of first order orientation-specific filters in the visual system, whose outputs are then transformed through a second stage non-linearity. A final stage of linear filtering is then conducted on this output with larger receptive fields than those of the first, and is sensitive to the location of any spatial gradient in the orientation-defined texture (see Bergen, 1991, for a detailed review of such a model).

Despite quite substantial and collective efforts, it has been difficult to isolate the cortical regions that are involved in this process. One candidate area is the earliest possible cortical stage of visual processing- primary visual cortex (V1). This is seen as a worthy substrate for at least the linear filtering stage of texture segmentation, due in part to its extensive network of simple and complex cells, and the presence of both excitatory and inhibitory lateral interactions, which may be sufficient to perform even 2nd-order non-linearity (Landy & Graham, 2004). Indeed, single-cell recordings in the macaque revealed the activity of V1 cells in response to an oriented line or texture to be enhanced by the presence of an orthogonal surround (Knierim & van Essen, 1992; Sillito, Grieve, Jones, Cudeiro & Davis, 1995), whereas activity was suppressed for homogenous regions (Nothdurft, Gallant & van Essen, 2000). Importantly, some researchers have suggested that V1 is sufficient in this role; when extrastriate areas are ablated or suppressed, cell responses in V1 to texture borders do not diminish (Lamme, Supèr, & Spekreijse, 1998; Lamme, Rodriguez-Rodriguez & Spekreise*,* 1999; Hupé, James, Girard & Bullier, 2001). Sensitivity to contrast-defined contours (another type of 2nd-order contour) has also been shown in area 18 of the cat (one of the first feline cortical processing stage along with area 17; Mareschal & Baker, 1998), and in humans, EEG source-imaging also points to activity in V1 being specific to the segmentation of a figure defined by an orientation contrast (Appelbaum, Ales & Norcia, 2012). Primary visual cortex, therefore, may play a sufficient role in texture segmentation.

This is not an entirely supported view, however; neuroimaging evidence in particular often reveals the blood-oxygen-level-dependent (BOLD) response in early visual areas (specifically V1 and V2) either not to be correlated with the degree of contrast of a texture-defined figure or to be much weaker than that in later areas (Kastner, de Weerd and Ungerleider, 2000; Larsson, Landy & Heeger, 2006[1]; Montaser-Kouhsari, Landy, Heeger & Larsson, 2007; Thielscher, Kolle, Neumann, Spitzer & Gron, 2008). Typically, the activity is observed in an area described as human V4, and some have directly speculated that the segmentation process depends on recurrent activity between this area and lower-level areas, with the final representation of the segmentation occurring in V4 (Thielscher *et al*, 2008). Computational neuroscience has attempted to further our understanding of this process, specifically by modelling V1, V2 and V4 cells together in an interactive network such that selective artificial "lesions" can be induced and the resulting activity in these regions can be observed in response to stimuli that require 2nd-order segmentation (Thielscher and Neumann, 2007). With the removal of feedback connections from V4 to V2, the neural representation of a region defined by an orientation contrast in the presence of orientation noise is suppressed at the level of V2 (Thielscher & Neumann, 2005), further supporting the view that V4 is necessary for segmentation. Studies focussing on the impairments following human lesions would be invaluable in determining if there is a single cortical area responsible for texture segmentation, but their numbers are sparse and neural damage is typically non-specific. One neuropsychology patient, with damage to V4, is impaired (although not completely) in segmenting a single bar defined by an orientation contrast to its background (Allen, Humphreys, Colin & Neumann, 2009). The patient, however, also has damage to the earlier areas V2 and V3 and so it remains unclear whether texture segmentation is indeed a process which necessitates the area identified as V4 in the human cortex.

It is important to note at this point that, although determining the true human homologue of macaque area V4 is difficult, the reference to human V4 here is made with

---

[1] Larsson *et al* (2006) reported only a trend for progressively greater activity from V1 to V4.

the intention to relate to the terminology of others who have conducted related studies of texture segmentation with human observers (e.g. Thielscher *et al*, 2008; Thielscher & Neumann, 2007; Allen *et al*, 2009; Kastner *et al*, 2000). The region in question is comparable on anatomical grounds to that identified by McKeefry and Zeki (1997), being situated on the collateral sulcus of the fusiform gyrus[2].

The purpose of the present study was to develop our understanding of the role of area V4 in texture segmentation, by testing the performance of a single patient (MS), who has bilateral damage to his ventromedial occipitotemporal cortices that includes the area often identified as human V4. Damage to his early visual cortex (at least in one hemisphere) is relatively spared. MS, therefore, provides a good neuropsychological test of the role of human V4 in texture segmentation. In the present study, patient MS performed an oddity-detection task, to indicate which of three texture patches contained a target subregion defined by an orientation contrast. His *discrimination* of the same stimuli was also tested, by requiring him to perform the task with black border masks at the locations of contrast, thus attenuating any perception of segmentation. If MS's intact brain regions allow him to segment the target, then his performance should be consistently greater when the border masks are absent than when they are present. In experiments one and two, the degree of orientation flow in the stimuli (i.e. by what degree the individual orientations were rotated clockwise as one moved rightwards and downwards through each texture pattern) was manipulated in order to vary the degree of orientation noise. With this manipulation, MS may have revealed a selective deficit in segmenting only when a considerable amount of orientation noise was present.

---

[2] Area V4 in the monkey was initially declared to be a "colour area" (Zeki, 1973, 1983, 1990; Walsh, Butler, Carden & Kulikowski, 1993) and the observation in humans that an area in the vicinity of the lingual and fusiform gyri responded selectively to chromatic stimuli (Zeki *et al* 1991), and that damage within this area resulted in achromatopsia (Heywood, Cowey & Newcombe, 1991), lead to the classification of this particular area as human V4. Surgical lesions of macaque V4, however, did not reveal a comparable animal model of achromatopsia, instead producing significant deficits in form percpetion (Heywood, Gadotti & Cowey, 1992; Heywood & Cowey, 1987), casting doubt on its colour selectivity and thus on its putative relation with human V4.

3.3. General methods

3.3.1. Subject

Patient MS has bilateral damage to his ventromedial occipitotemporal cortex and is profoundly achromatopsic and prosopagnosic as well as having visual object agnosia. Damage to the left hemisphere includes the temporal pole, parahippocampal and fourth temporal gyri of the temporal lobe, the collateral sulcus, and the mesial occipitotemporal junction. His achromatopsia is most likely explained by the damage to the lingual gyrus and anterior collateral sulcus, typically associated with “colour area” V4. Damage in this vicinity is also likely to be the cause of his impairment in discriminating surface properties of objects more generally (Cavina-Pratesi, Kentridge, Heywood & Milner, 2010b) along with his prosopagnosia. His ability to perceive object form remains intact, which is likely the result of his spared lateral occipital complex (LOC; Cavina-Pratesi *et al,* 2010b). The damage in his right hemisphere is more extensive than that in the left, including primary visual cortex resulting in a homonymous hemianopia with macular sparing, whereas his occipital lobe in his left hemisphere is largely intact. For a more extensive case description of MS, see Heywood, Cowey and Newcombe (1994). Testing was conducted over the course of two days at the University of Durham’s Psychology Department. MS was 63 years of age at the time of testing.

3.3.2. Stimuli

MS viewed the display monitor at a distance of roughly 80cm. Stimuli were presented on the uniform grey background (50 cdm^{-2}) of a ViewSonic 17’’ (1254 x 877 pixels) colour monitor driven by a Cambridge Research Systems VSG 2/5 Graphics System running at 100Hz.

Stimuli consisted of three separate textures of 10 x 10 Gabor patches (Michelson contrast of 90%). Each Gabor patch measured 0.7° in diameter, with a spatial frequency of 1.4 cyc/° and an envelope with a standard deviation of 0.2°. Each Gabor patch was separated from its neighbours by 0.3°. Thus, each texture measured 9.7° in width and in height. The phase of each Gabor was determined randomly and independently from the full 360° cycle. The three patches were aligned vertically in the centre of the display, with a distance of 1.4° separating each of the top and bottom patches from the middle. See **figure**

3.1 for an illustration of the stimuli. MS was allowed full free-viewing of the stimuli in order to account for his left hemianopia.

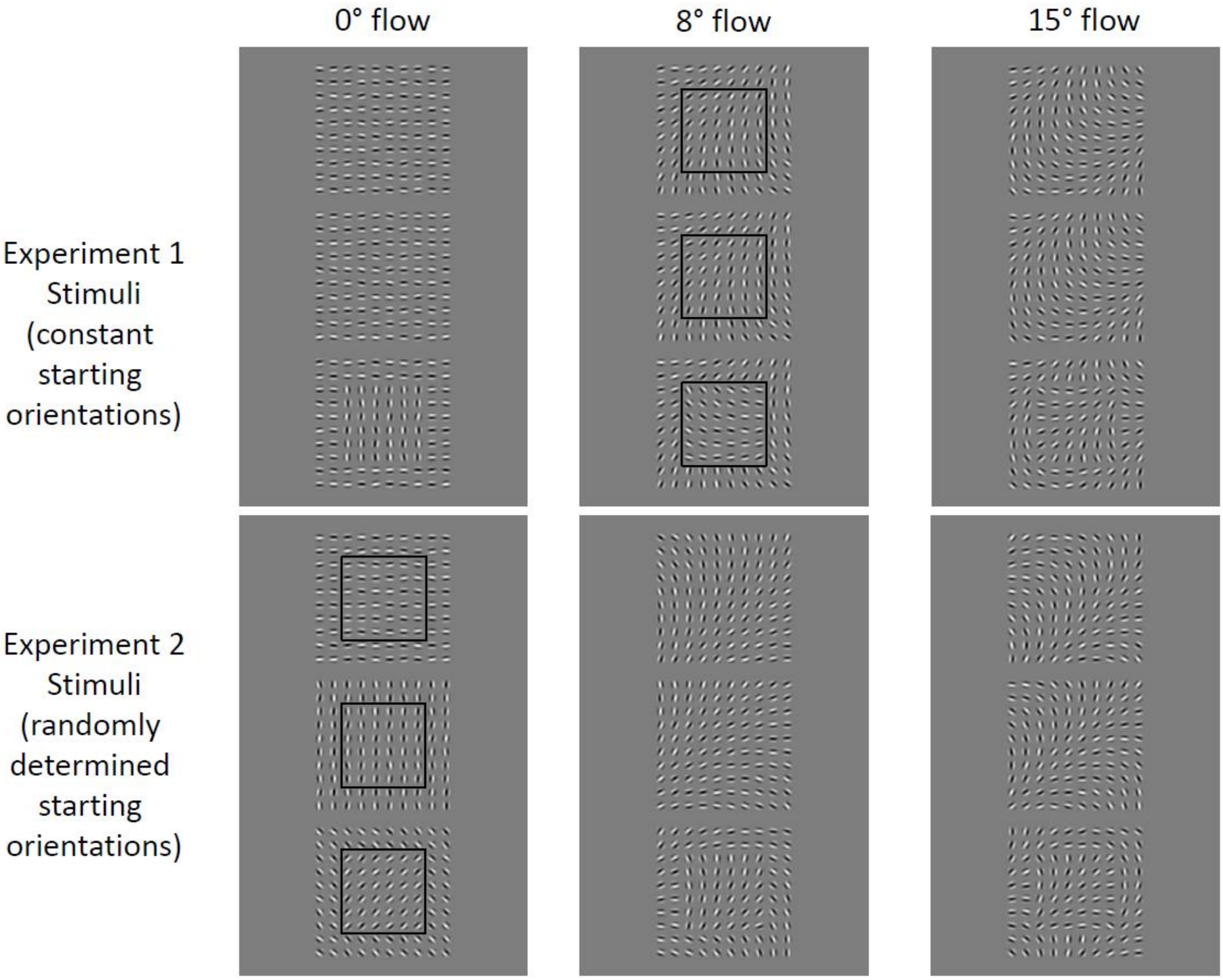

Figure 3.1 Illustration of the stimuli used in the the first two experiments (top row: experiment 1; bottom row: experiment 2). From left to right, the stimuli increase in the degree of orientation flow. In experiment 1, the starting value of the orientation flow was constant (i.e. the top-left Gabor patch set at 45°) and the same for the three textures. In experiment 2, starting values were randomly allocated to the three textures (each 45° different to the other two), thus removing the possibility of completing the task on a purely local scale. Two of the images show examples of the discrimination condition, in which the segmentation of the target is masked by black outlines. Accurate performance in this condition requires more effortful analysis of the target and background regions. If MS is unable to segment, however, then his performance should be similar with and without the border masks. MS's task was to detect which texture (top or bottom only) was the odd-one-out.

3.4. Experiment 1 – 2nd-order segmentation in patient MS

3.4.1. Methods

Three levels of task difficult were included and randomised throughout each block of trials; the difficulty was determined by the degree of orientation flow within the three textures. That is, the degree to which the orientation of each Gabor patch was tilted relative to those immediately above and to the left, originating at the uppermost and leftmost Gabor patch (which was set at vertical). Thus the orientation of each Gabor patch (*g*) at location (*i j*) is represented in the following function:

$$g_{(i,j)} = a + b(i - 1) + b(j - 1)$$

Where *a* represents the starting orientation, *b* represents the degree of orientation flow and *i* and *j* represent the horizontal and vertical positions of the Gabor, respectively. For this experiment, all three textures began with a vertical orientation (a=0), and three levels of orientation flow were used: b=0° (no flow), 8° (small flow) or 15° (large flow). In addition, each individual Gabor was drawn from a Gaussian distribution with a standard deviation of 5°, thus introducing a small amount of baseline orientation noise, and this was done *de novo* on each trial.

In each trial, one of the textures (either the top or bottom, chosen randomly with equal probability) was constructed differently to the others, by offsetting the central 6 x 6 square region of Gabor patches by 90°. In half of the blocks of trials, the central 6 x 6 region in all three textures was surrounded by a black border mask 0.2° in thickness, thus attenuating the perception of the target's contours. Patient MS conducted ten blocks of trials: five with the border masks absent (segmentation condition) and 5 with the border masks present (discrimination condition). The two block types were conducted in alternation, and within each block, 10 repetitions of each of the three flow pattern levels were randomly interleaved, amounting to 30 trials per block. The stimuli remained onscreen until MS indicated which patch (the top or bottom) appeared to be the odd-one-out (i.e. which contained the central 6 x 6 target). Patient MS verbalised his response and the experimenter pressed the corresponding key on a response box to proceed to the next trial. Examples of the types of stimuli are shown in **figure 3.1**.

3.4.2. Results

For each level of orientation flow (0°, 8° and 15° respectively), MS's segmentation accuracy was significantly above chance: 50/50 (100%; $p<0.001$), 46/50 (92%; $p<0.001$) and 38/50 (76%; $p<0.001$). Similarly, his discrimination accuracy was also above chance: 47/50 (94%; $p<0.001$), 32/50 (68%; $p=0.0325$) and 32/50 (68%; $p=0.0325$). Accuracy was collapsed across the three levels, and overall MS performed better in the segmentation condition than in the discrimination condition $\chi^2_{(1)} = 11.78$, $p<0.001$. Results are illustrated in **figure 3.2a**.

3.5. Interim discussion

MS's discrimination ability in the previous experiment, although less accurate than his segmentation ability, was nonetheless above chance. It could be argued, however, that MS was completing this task on a very local scale, by attending to and comparing only the top-left orientation, for example, within the 6 x 6 target region of each of the three textures. Experiment 2, therefore, was designed to remove this local cue, and thus to test more appropriately MS's discrimination of the regions. This was done by assigning each of the three textures with a unique starting orientation, thus ensuring that it was not possible to infer the location of the target region from one single orientation within the target region.

3.6. Experiment 2 – 2nd-order segmentation in patient MS without local orientation cues

3.6.1. Methods

The stimuli and methods were identical to experiment 1, except that for each trial a random value from the full 360° cycle was chosen and was assigned as the starting orientation for one of the textures. Of the remaining two textures, one was chosen to have a starting value 45° clockwise from this value, and the other 90°, and the positions of these three textures were randomly allocated in the vertical array. Thus, each texture assumed different background orientations to the other two, and the odd-one-out corresponded to the one whose target region did not agree with the continuity of the background. This task could not be accomplished on a local scale.

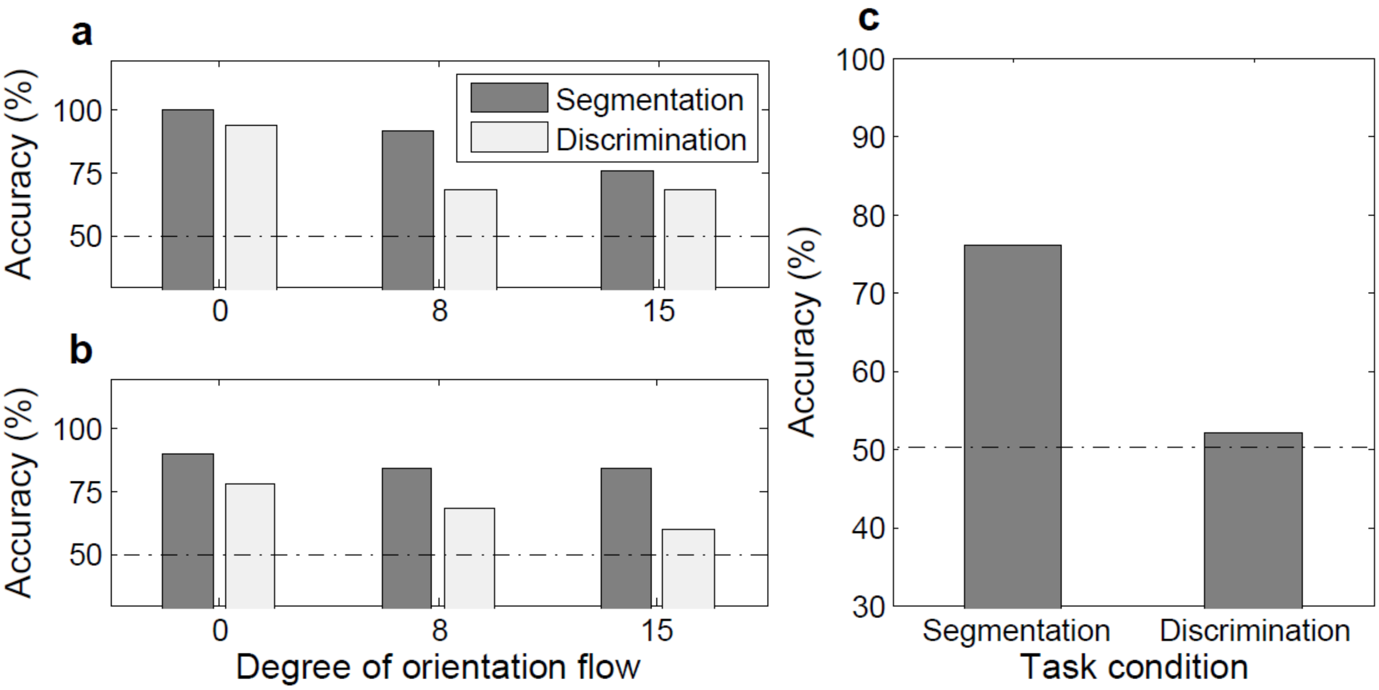

Figure 3.2 MS's accuracy (%) in experiments 1 (a), 2 (b) and 3 (c). In experiments 1 and 2, MS was able to segment the target more accurately than he could discriminate (i.e. when black borders masked the target contours). In experiment 3, when the orientations of the target and background regions alternated at a frequency of 12.5Hz, MS was still able to segment (76% correct), whereas his ability to discriminate the regions was at chance (52%). Together, these results indicate that MS has an intact ability to segment a target defined by an orientation contrast, and that this ability is similar to that of normal observers.

3.6.2. Results

For each level of orientation flow (0°, 8° and 15° respectively), MS's segmentation accuracy was significantly above chance: 45/50 (90%; $p<0.001$), 42/50 (84%; $p<0.001$) and 42/50 (84%; $p<0.001$). His discrimination accuracy was also above chance, but not in the most difficult condition: 38/50 (76%; $p<0.001$), 34/50 (68%; $p=0.008$) and 30/50 (60%; $p=0.101$). Accuracy was collapsed across the three levels, and overall MS again performed better in the segmentation condition than in the discrimination condition $\chi^2_{(1)} = 13.72$, $p<0.001$. Results are illustrated in **figure 3.2b**.

3.7. Interim discussion

In the previous experiments, MS was consistently more accurate when segmenting than when discriminating the textures. This implies that when the orientation contrast was present (i.e. not masked by black outlines) MS was using that to inform his judgment and was not relying solely on a technique of regional discrimination. This was also true for stimuli in which there was a considerable degree of orientation noise, suggesting that, contrary to what has been implied through computational modelling (Thielscher & Neumann, 2003), V4 is not required for segmentation of even complex 2nd-order stimuli.

Experiments one and two, however, did not directly address an important aspect of segmentation, namely that it is rapid and potentially phase-insensitive (Norman, Heywood & Kentridge*,* 2011). The mammalian visual system gives temporal priority to segmentation, relative to regional discrimination, as it shows an enhanced cellular response corresponding to the location of a border prior to the location of a texture region (Lamme *et al,* 1999; Romani, Caputo, Callieco, Schintone and Cosi, 1999). Many psychophysical experiments in humans also lend weight to the idea of a temporal dissociation between segmentation and discrimination (e.g. Motoyoshi, 1999), and together imply that very early, rapid mechanisms underlie segmentation. In experiment 3, therefore, MS's segmentation ability was tested using orientation-reversing stimuli (Norman *et al,* 2011), in which the target and background regions alternated at a frequency of 12.5Hz. At approximately this frequency level in normal observers, the ability to segment persists but the ability to discriminate is likely to be lost or severely impaired. Demonstrating that MS could successfully segment these stimuli, whilst being unable to discriminate the regions he was segmenting, would substantially strengthen the claim that MS can segment just like normal observers.

3.8. Experiment 3 – phase-insensitive 2nd-order segmentation in patient MS

3.8.1. Methods

Stimuli were constructed in the same way as in experiment 1, although the level of orientation flow was not manipulated (set at 0°). Stimuli were presented onscreen following a brief mask (30ms) consisting of three textures of random orientations. The orientations of the test stimuli then reversed at a frequency of 12.5Hz between vertical and horizontal (40ms per stimulus frame). See **figure 3.3** for a depiction of the stimuli. This persisted until MS made a response. Four blocks were conducted; two of which included border masks and two of which did not, and each block contained 25 trials.

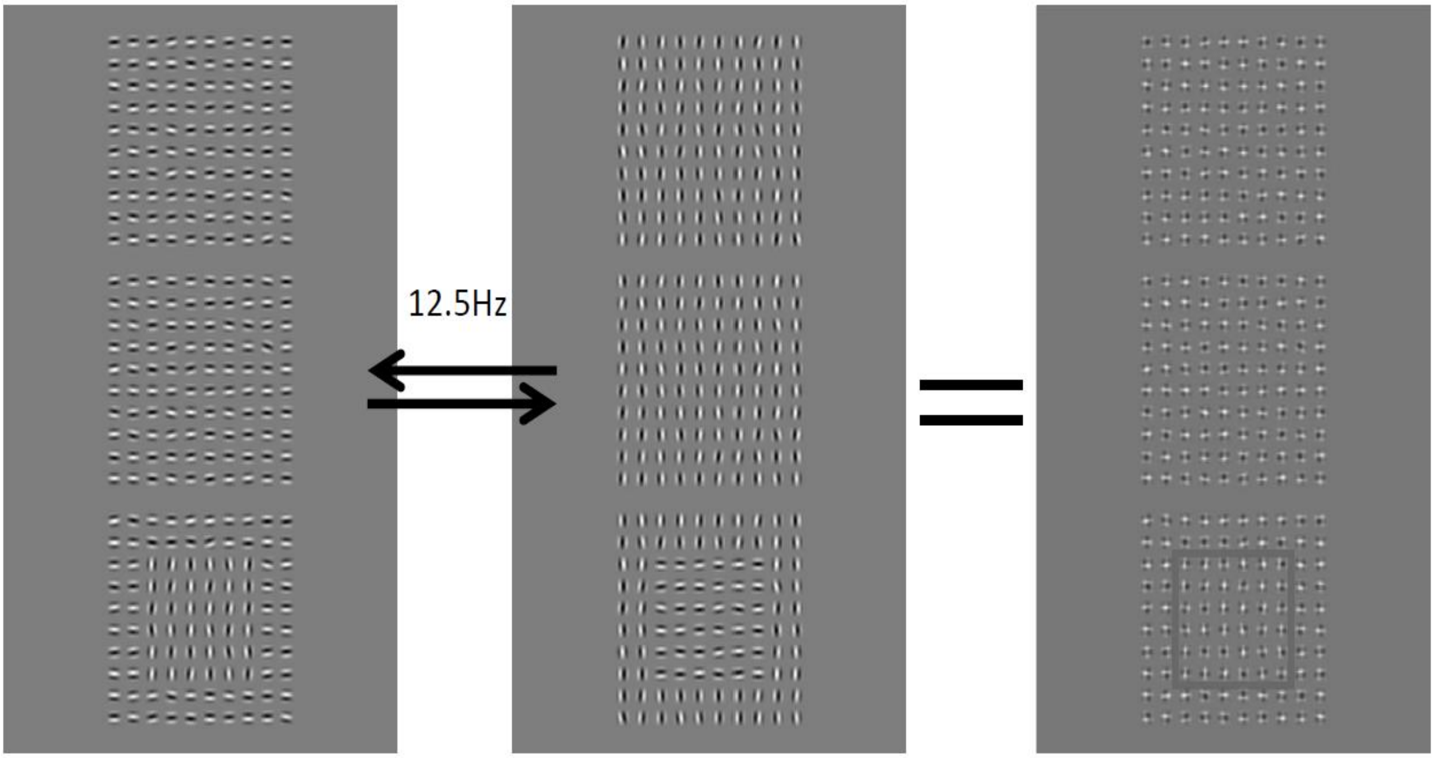

Figure 3.3 Illustration of the orientation-reversing procedure used in experiment 3. The target is defined by an orientation contrast of 90°, but the phase of this contrast is reversed at a frequency of 12.5Hz. The resultant perception is illustrated in the rightmost illustration panel; the target and background regions appear identical yet the perception of an "illusory" contour is present around the target (as a result of the rapid segmentation process that may be phase-insensitive). MS's task, again, was to detect which texture was the odd-one-out. He also completed a version in which black border masks (see **figure 3.1** for an example) were present at the location of the contrast to test his discrimination of the regions.

3.8.2. Results

MS performed the segmentation task significantly above chance, with 38/50 hits (76%; $p<0.001$), which is in stark contrast to his performance on the discrimination task, which remained at chance (26/50 hits, 52%; $p=0.444$), with his performance being statistically greater in the segmentation condition ($\chi^2_{(1)} = 6.25$, $p=0.012$). Results are illustrated in **figure 3.2c**.

3.9. General discussion of study two

If MS's damaged cortex were responsible for texture segmentation, then he should be unable to detect a region defined by an orientation contrast. In experiments one and two this was clearly shown not to be the case; MS achieved close to 100% accuracy in the simplest segmentation task (no orientation flow). Even when a large degree of orientation flow was introduced to the stimuli, which requires more complex analyses (Thielscher & Neumann, 2007), MS's accuracy in the segmentation condition did not fall below 75%. These results strengthen the claims that human area V4 is not necessary for texture segmentation and that early cortical areas are likely to be sufficient.

Is it possible, however, that MS's perception of the segmenting region was not automatic and immediate, and instead his performance relied on a more effortful discrimination process? This is important to establish, as true segmentation must be rapidly achieved and be effortless. The results of experiments one and two, however, do suggest that he was indeed segmenting; when the contour of the segmenting region was masked by a black border, MS performed the task with significantly less accuracy, albeit still above chance in most cases. This indicates that his perception of the target in the segmentation conditions was aided by the contrast information present at the location of the contours. It cannot be claimed, therefore, that he was using a form of regional discrimination (i.e. scrutinising individual orientations of target and background) in order discern the location of the target in the segmentation conditions, as his performance would have been unaffected by the presence of the border masks if this were the case.

Furthermore, experiment three strongly suggests that MS's segmentation is rapid and like that of normal observers. MS was able to identify the target even when the orientations of the background and target elements alternated at a frequency of 12.5Hz (40ms per stimulus frame). In normal observers at approximately this frequency, the background and target regions become indistinguishable from one another, yet the perception of a contour between the two is strong and segmentation still occurs (Norman *et al*, 2011). MS showed this same pattern of results; he was at chance when discriminating the regions (i.e. when the target contour was masked), yet he could accurately use the contour information when it was present. Thus, despite the strict temporal limit imposed upon the processing of the stimuli, MS was nonetheless able to segment. It is very unlikely, therefore, that MS was relying on a method other than segmentation to achieve this task.

These results shed new light on the neural processes underlying texture segmentation. Specifically, they provide strong evidence that texture segmentation (by orientation contrast at least) does not necessitate the areas that are damaged in MS, specifically what is considered area V4. This contradicts earlier evidence from neuroimaging studies (Thielscher *et al,* 2008; Kastner *et al,* 2000) and computational modelling (Thielscher *et al,* 2007), which have led to the theory that V4 is necessary for segmenting figures defined by an orientation contrast, with specific emphasis on its role in providing feedback to lower areas and in ultimately representing the outcome of the segmentation process. The new results do, however, corroborate evidence from earlier cell-recordings in non-human studies, which indicated primary visual cortex to be sufficient for segmentation (e.g. Lamme *et al,* 1998; 1999; Hupé *et al,* 2001). Recently, Hallum, Landy and Heeger (2011) provided evidence from functional magnetic resonance adaptation showing populations of neurons within V1 to be sensitive to 2nd-order contrast, with the observed activity reflecting a two-stage filter process with surround suppression from 1st-order filters. As they found comparable stimulus-specific adaptation throughout V1-V4, they conclude that this activity was the result of feed-forward activity from V1.

Cells within V1 are known to process chromatic contrast (Kentridge, Heywood & Weiskrantz, 2007), which are highly likely to underlie MS's preserved ability to perceive form on the basis of such information whilst being blind to colour surface identity (Heywood *et al*, 1991), with regional colour information being assigned at a later cortical stage (within the lingual gyrus and anterior collateral sulcus; Cavina-Pratesi *et al*, 2010b). As shown in the present study, however, MS retains a generally good ability to discriminate regions of orientations when the borders are masked; his perception of the stimuli, therefore, is not entirely analagous to that of colour. His ability to discriminate regions of orientations above chance in most cases is likely to be mediated by the orientation-selective cells within V1.

Primary visual cortex is, however, not unique in its apparent sensitivity to texture contrast, and from the current results it is not possible to truly conclude that it is solely primary visual cortex that is responsible for MS's segmentation ability. Specifically, the second cortical visual area (V2), which lies immediately adjacent to V1 and with which it is largely interconnected, may play an important role in segmentation and it is possible that at least part of this area may be preserved and functional in MS's left hemisphere. Schira, Fahle, Donner, Kraft and Brandt (2004) showed that the BOLD activity in V2, although not V1, correlated with the contour of a figure defined by an orientation contrast, and they hypothesised that the observed activity in mid- to high-level areas merely reflected the feedforward activity from V2. Importantly, in their study, subjects completed an attentionally demanding foveal letter discrimination task and as a result the segmenting stimuli did not reach awareness. The activity they found, therefore, is likely to reflect the preattentive processes involved in segmentation. Schira *et al* also focussed their analysis only on the retinotopic region associated with the texture *contour*, enabling them to more accurately relate the observed activity to the processes of segmentation rather than the more general processes associated with figure perception. V2 cells, both at the individual and population level, also respond well to complex shape information (Hegdé & Van Essen, 2000, 2003) and in a way that evolves relatively quickly (as early as 40ms of stimulus exposure; Hegdé & Van Essen, 2004). This is largely upheld by the finding that lesions of V2 produce significant impairments in segmenting a texture-defined figure (Merigan, Nealey & Maunsell, 1993), in which the orientation of a figure comprising a subset of line segments of different orientations to their background could not be discriminated, whereas the same figure defined by colour or size could be.

It is also important to address the role of LOC in the perception of a target defined by 2nd-order segmentation, which has been identified as a potential correlate, along with V4, in some neuroimaging experiments (e.g. Thielscher *et al,* 2008; Larsson *et al*, 2006). The role of this area has been extensively studied and seems to represent a neural correlate of the perception of object shape seemingly regardless of how the shape is defined (e.g. Kourtzi & Kanwisher, 2001; Tyler, Likova & Wade, 2004). In the present study, however, MS was not required to make any discrimination of the form of the target; the task required only its localisation. It is likely, therefore, that this task could be achieved independently of the LOC, and that the activity observed in this region in neuroimaging studies is not directly a result of the segmentation process *per se*, but instead reflects that of an integration stage for the purposes of form perception beyond the early- to mid-cortical processing hierarchy (Thielscher *et al,* 2008). This is a notion that may be indirectly supported by monkey lesions;

Merigan (2000) lesioned area V4 in monkeys, an area which shows many processing characteristics not unlike those of LOC in humans and is thus not homologous to the human V4 discussed here, and demonstrated that the ability to segment two abutting regions of orientation orthogonal to one another was preserved. The same animals could not, however, identify the orientation of a bar defined by the same orientation contrast. This may suggest that their segmentation mechanism was intact, but their ability to integrate this information to perceive form was not.

In conclusion, the results from these experiments with patient MS provide compelling evidence that human V4 is not necessary for texture segmentation. Populations of neurons within primary visual cortex may be sufficient, although it is not known yet exactly how they might achieve this.

4. Study Three – The early encoding of orientation variance in the visual system

4.1. Abstract

Our discrimination of regional irregularity, specifically of orientation variance, seems effortless when we view two patches of texture which differ in this attribute. Spatial gradients in orientation variance, however, do not automatically segment, unlike equivalent gradients in mean orientation. Little is understood of how the visual system might process this second moment statistic of orientation, but there is some evidence to suggest that it is explicitly encoded by the visual system by populations of neurons tuned broadly to high or low levels of variance. This theory is supported in the present study, which shows that selective adaptation to low or high levels of variance results in a perceptual aftereffect that shifts the perceived level of variance of a subsequently viewed texture in the direction away from that of the adapting stimulus (experiments 1 and 2). Importantly, the effect is durable across changes in the mean orientation of the elements that constitute the stimuli, suggesting the existence of a variance-sensitive component of perception that is independent both to local orientation components and to global first moment orientation statistics (i.e. mean orientation). Some research has shown that as visual encoding becomes increasingly complex and involves higher areas in the visual cortex, it also increases in its spatiotopy, which can be revealed through spatiotopic aftereffects (provided that the stimulus is attended). It was therefore predicted that, in experiment 3, the variance-specific aftereffect might show signs of spatiotopic encoding, unlike the equivalent aftereffect of adaptation to the first moment orientation statistic (the tilt aftereffect), which is represented in primary visual cortex and exists only in retinotopic coordinates. The variance aftereffect is shown in this study, however, also to be purely retinotopic. Finally, it is shown in experiment 4 that patient MS, who has intact early cortical areas but not mid-level ventral areas, retains sensitivity to orientation variance. Together, these results suggest that region-based property of orientation variance is encoded explicitly by the visual system and at an earlier cortical stage than previously theorised.

4.2. Introduction

Visual segmentation is achieved by extracting discontinuities in 1st- and 2nd-order visual attributes. This aspect of visual perception is *edge*-based, in that information processing begins at the locations of contours and spreads “inwards”. This is not the only aspect of visual texture perception, however; certain spatial gradients require *region*-based discrimination that precludes automatic segmentation. In study one of this thesis, for instance, it was shown that spatial gradients in orientation variance do not rapidly segment. In fact, discrimination performance was more rapid when the patches were separated relative to when they were abutting – the complete opposite effect of edge-based perception. How we are capable of representing and perceiving *regional* texture properties of such nature is not clear and, although our understanding of texture segmentation processes is far from complete, we have a considerably better grasp on those processes than we do in understanding how the visual system encodes region-based properties of an image.

Summary statistics, which offer quick estimations of various image properties, are thought to play at least a partial role in some forms of region-based perception. Summary statistics are available in many different forms, and human performance on certain tasks frequently correlates with predictions from models based on statistical summations. Ariely (2001) demonstrated that when a test circle was presented following a set of similar circles of varying radii, observers could determine whether or not the radius of the test circle was equal to that of the mean of the set, despite not being able to reliably identify whether the test circle was actually contained in that set. This represents a dissociation between the perception of a summary representation that describes a set and that of a single unit within that set, suggesting the workings of a dedicated mechanism that extracts region-based information that is statistical in nature. Chong and Treisman (2003) reported similar results, and found that when they varied a number of properties of the sets (e.g. density, numerosity and average size) only manipulations of average size affected perceived size of the circles. The ability of the visual system to reliably encode statistical properties is apparent elsewhere; for instance in estimating the mean direction among individual local motion paths (Williams & Sekuler, 1984) and in estimating the mean orientation amongst oriented line segments (Dakin, 2001; Dakin & Watt, 1997). The visual system relies on such summary estimations in order to be economical, minimising computational effort and preserving just enough information when confronted with detailed environments (Ariely, 2001; Morgan, Mareschal, Chubb & Solomon, 2012). Inherent in this notion is the assumption that most surfaces and regions in the environment are uniform with few, if any, discontinuities.

There are more complex region-based characteristics, however, which require statistical estimation of a higher order. The encoding of orientation variance (a second moment statistic[3]), for instance, may represent one process by which regional irregularity on object surfaces is perceived, which could inform object recognition. Orientation variance is of particular interest because the visual system, by its very nature, is a noisy environment, perturbing much of the information that it receives. In fact, estimations of a single orientation are sampled from a Gaussian distribution with a slight standard deviation centred round the actual orientation. Morgan and colleagues (Morgan, Chubb & Solomon, 2008; Morgan *et al,* 2012) point out the apparent contradiction between this attribute of the visual system and our perception of homogenous texture: in a field of perfectly uniformly elements we should see deviations from the mean, yet this is not the case. They theorise that in instances of textural analyses like this, our perception is determined not only by our internal representation of individual orientations, but also by a dedicated mechanism that explicitly estimates the variance. They provide evidence for this in finding a "dipper function" for discriminations based on orientation variance of the type that has also been documented in discriminations of blur (Watt & Morgan, 1983) and contrast (Nachmias & Sansbury, 1974). This threshold nonlinearity is thought to allow the visual system to discount its own imperfections (intrinsic noise) when conducting a statistical estimation, and only when the estimation of variance exceeds a certain criterion are the individual deviations apparent.

Experiments one and two here explore whether orientation variance is encoded as a property of a texture, using selective adaptation. Adaptation describes the prolonged sensory exposure to a particular stimulus, and often results in perceptual aftereffects that result in a shift in perception *away* from the attributes of the adapted stimulus. Adapting to

---

[3] The term "second moment" statistic is used to describe orientation variance as being of a higher order than the mean (first moment) but lower than the kurtosis or skewness (third moment), not to be confused with the term "2nd-order" stimuli.

upward motion, for instance, will result in the perception of downward motion of a subsequently viewed image that is actually static. The dominant theory of adaptation is that the perception of the adapted attribute is determined by the balance of activity between populations of neurons tuned to different levels of that attribute, and that the neurons that are firing during exposure to the adapted stimulus become *fatigued* or *inhibited (*Krekelberg, Boynton & Van Wezel, 2006). When a subsequent stimulus is experienced, the drop in sensitivity of these neurons relative to those that are non-adapted results in an imbalance that creates a “misperception” of the adapted attributes in the direction away from the adapted level.

Adaptation has previously been used as a tool to explore other fundamentals of region-based perception. Corbett, Wurnitsch, Schwartz and Whitney (2012), for instance, recently showed selective adaptation aftereffects for the mean size of a set of circles, corroborating earlier evidence that suggested a dedicated, and independent mechanism, for the extraction of mean size in an image. Additionally, contrast in an image, which has been likened to extracting the standard deviation (or variance) of luminance information (i.e. a second moment statistic), and density, which has been likened to extracting the same statistics from contrast itself (i.e. luminance *kurtosis*; a third moment statistic), are both properties that can be selectively adapted in the visual system (Durgin, 2001). Adapting to images with positively and negatively skewed luminance distributions has been shown to produce perceptual aftereffects in the opposing direction, which manifests as an increase or decrease in the perception of gloss on a surface (Motoyoshi, Nishida, Sharan & Adelson, 2007; although see Kim & Anderson, 2010). These are examples of selective adaptation revealing dedicated mechanisms in the visual system that are sensitive to specific statistical properties of images. Importantly, such effects are measurable on surfaces that resemble *real* materials and objects, and are not restricted to abstract psychophysical displays (Durgin & Huk, 1997; Motoyoshi *et al*, 2007).

In experiment one, subjects were continually presented with two textures during an adaptation phase; one on the right and one on the left. One of these textures was always constructed from a Gaussian distribution of a *medium* variance, and the other from either a relatively *low* or *high* variance. When presented with a similar pair of textures in the test phase, subjects were required to identify which texture (left/right) appeared to contain the most variance and the level of perceived subjective equality (PSE) was sought. Adapting to a low variance on one side should *increase* the perception of variance of the test texture on that side, and *vice-versa* for adaptation to high variance. Importantly, mean orientations of the adapting and test stimuli were randomly determined for each presentation to negate low-level effects of orientation adaptation.

4.3. Experiment 1 – orientation variance adaptation with randomised means

4.3.1. Methods

Participants

8 observers took part in the study. Participants gave their written, informed consent.

Equipment and stimuli

The display monitor was viewed at a distance of 41cm (subjects rested their head on a chin rest). Stimuli were presented on the uniform grey background (50 cdm^{-2}) of a ViewSonic 17'' (1254 x 877 pixels) colour monitor driven by a Cambridge Research Systems VSG 2/5 Graphics System running at 100Hz.

The experimental stimuli used both in the adaptation and test sections consisted of a pair of textures, each comprising 9 x 9 Gabor patches. Each Gabor patch measured 0.6° in diameter and was separated from its neighbours by 0.3° with a small positional jitter. Each texture, therefore, measured 7.8° in width and in height. All Gabor patches had a spatial frequency of 0.6 cycles/degree and each was assigned an individually randomly determined phase from the full 360° cycle. The Michelson contrast of each Gabor was 90%. The pair of textures was presented in the centre of the screen for both the adaptation and test phase.

Each texture's orientation statistics were determined by independently drawing each composite orientation value from a Gaussian distribution with a particular mean (μ) and standard deviation (σ). For both adaptation and test flashes, μ was randomly determined on each presentation. The value of σ for each texture depended on several experimental parameters; during adaptation, one of the textures assumed a medium level (σ=10) and the other assumed either a low (σ=2) or high (σ=26) level (to be known as the "adapting texture"). These values were chosen from a period of pilot testing on one of the participants as values which induced aftereffects that were comparable in magnitude to one another. Examples of the stimuli are shown in **figure 4.1**.

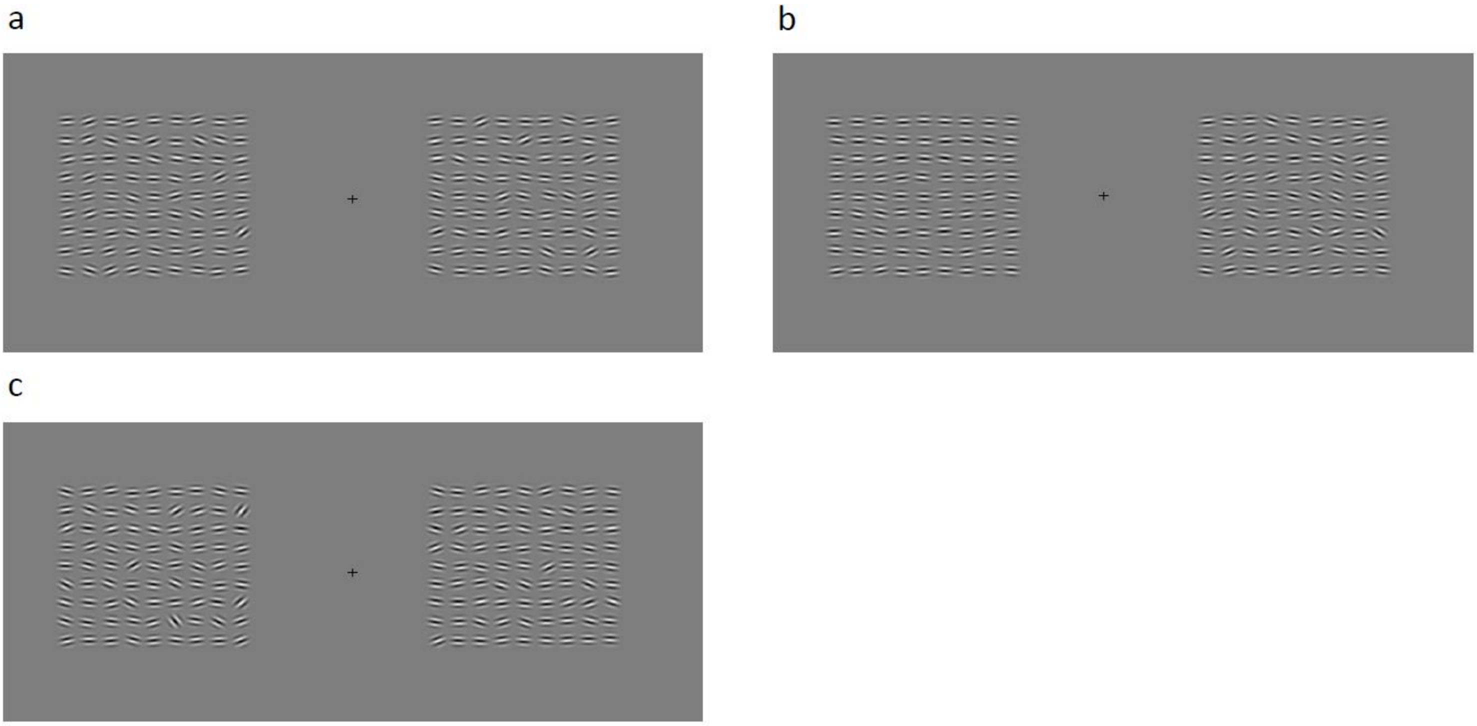

Figure 4.1. Examples of the different levels of variance used during adaptation in the experiments. In **a**, both textures are of a medium variance. In **b**, the left texture (in this case, the “adapting” texture) is of a low variance, and the right of a medium variance. In **c**, the left texture (again, the “adapting” texture) is of a high variance, and the right, again, of a medium variance.

Experimental conditions and design

Each subject completed two testing sessions, factoring the variance (low/high) of the adapting texture and half of the subjects received the adapting texture on the left and the other half on the right of fixation. Testing was conducted over a period of two days, with one session per day and the order of the sessions was counterbalanced across subjects.

Trial sequence

Subjects fixated a central cross. Each session began with a 60s adaptation procedure consisting of 30 stimulus flashes, each lasting 1s and separated by an interval of 1s. Each flash contained the medium variance texture and the “adapting texture”. For each adaptation flash, the textures were constructed *de novo*, with newly determined random mean orientations and each orientation drawn again from their relative distribution. Following adaptation, subjects received a warning tone and a pause of 750ms before the test stimuli flashed onscreen for 500ms. Subjects indicated which of the two textures in this flash appeared to contain the most orientation variance by pressing one of two keys, marking the end of one trial. Between trials, subjects were presented with a further top-up adaptation consisting of 3 flashes, and after every 25th trial, a prolonged top-up adaptation of 30 flashes.

In each session, trials were presented using two randomly interleaved one-up one-down staircases which measures the level of PSE. The variance of the test stimulus texture corresponding to the side of the adapting texture remained constant (at the medium level), and each staircase automatically adjusted the variance of the opposite texture until it reached an estimate of the subject's PSE (i.e. at which point they chose left/right in equal proportion). Two staircases were used; in one staircase, this texture began the session with a higher variance to the texture on the adapting side, and in the second staircase began with a lower variance, with a magnitude of 5 standard deviation (σ) units. The variance was incrementally increased/decreased initially by a magnitude of 2 σ units (first 4 trials per staircase) and from thereon 1 σ unit, until a criterion of 6 reversals had been met. Threshold estimation was acquired by averaging the final 4 reversal points.

4.3.2. Results

The measured threshold levels corresponded to the PSE (i.e. perceived level of variance of the test texture on the adapted side). This was done by systematically adjusting the variance of the non-adapted patch according to the subject's response until their discrimination of which texture was the most varied rested at 50%. This estimation was taken by averaging the results of two interleaved staircases for each condition. Data were averaged from all eight participants.

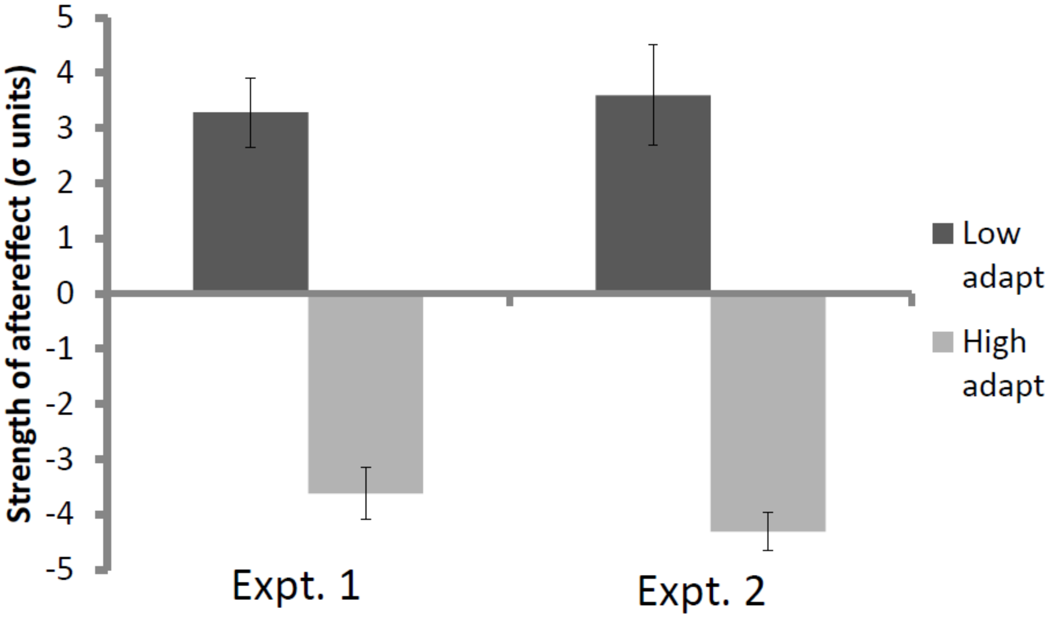

Figure 4.2. Results from experiments 1 and 2. Levels of PSE as a result of adapting to levels of low and high variance. A value of 0 corresponds to the actual variance of the texture (objective equality), as shown by the bold horizontal axis. Thus, a value above this level indicates a shift in variance perception towards the higher end of the scale and a value below indicates a shift towards the lower end. Error bars show +/- 1 SEM.

For the high-adaptation condition, the PSE was 3.6 σ units *below* the level of objective equality, whereas for the low-adaptation condition, the PSE was 3.2 σ units *above* the level of objective equality (see **figure 4.2**). Two one-sample t-tests were carried out between these PSE estimates and 0 (objective equality). Both were significant, with the measurement following high-adaptation being significantly lower than the objective measurement ($t_{(7)}=9.82, p<0.001$) and the one following low-adaptation being significantly higher ($t_{(7)}=4.92$, $p=0.002$). These results indicate that when observers adapted to a texture with low orientation variance, a subsequent texture appearing in the same location appeared to contain more variation than it actually did. The opposite was found following adaptation to high levels of variance.

4.4. Interim discussion

Experiment one has shown that the perception of orientation variance (a second moment orientation statistic) can be selectively altered by adaptation-induced aftereffects. Importantly, as the mean orientation was randomly allocated for each adaptation and test flash, the effect is not an indirect aftereffect following adaptation to low-level orientation signals. Additionally, the results are suggestive of a variance-sensitive mechanism that is independent to the low-level statistical properties of the texture region (i.e. mean orientation). This can not be determined conclusively from the first experiment only, however, because we cannot be certain that the mean orientation of the test textures on a given trial was not exposed at least once during adaptation, which may have been sufficient to produce a variance aftereffect that was specific to that mean orientation. Experiment two, therefore, replicated the first experiment but, instead of randomly determining the mean orientation on each flash, the adaptation textures were shown only with a mean orientation of 90° (corresponding to horizontal) and the test patches were only shown with a mean orientation of 0° (corresponding to vertical). If an aftereffect is reproduced in this experiment it would, therefore, strongly suggest that information that is encoded regarding orientation variance is general and is not restricted specifically to a particular mean orientation.

4.5. Experiment 2 – orientation variance adaptation across perpendicular means

4.5.1. Methods

See previous experiment.

4.5.2. Results

For the high-adaptation condition, the PSE was 4.3 σ units *below* the level of objective equality, whereas for the low-adaptation condition, the PSE was 3.6 σ units *above* the level of objective equality (see **figure 4.2**). Two single sample t-tests were carried out between these PSE estimates and 0 (objective equality). Both were significant, with the measurement following high-adaptation being significantly lower than the objective measurement ($t_{(7)}=12.50$, $p<0.001$) and the one following low-adaptation being significantly higher ($t_{(7)}=3.95$, $p=0.006$). These results thus replicate the findings from experiment 1 but, importantly, this occurred despite the adaptation textures constantly having a mean orientation that was 90° in contrast to that of the test textures.

4.6. Interim discussion

Together, experiments one and two have established that the perception of orientation variance (a second moment orientation statistic) can be selectively altered by adaptation-induced aftereffects. Importantly, the effects could only have emerged if the variance was encoded at a level that was independent of the local orientations and the first moment statistic of orientation. This is highly suggestive of a dedicated mechanism within the visual system that extracts the statistical irregularity amongst orientation signals in a given texture, agreeing with work by Morgan and colleagues (2008; 2012) and other work more generally that has found evidence through adaptation aftereffects of statistical computations by the visual system (e.g. Corbett *et al*, 2012).

Experiment three aimed to uncover the reference frame in which this variance aftereffect is based. The initial input into the visual system is based strictly in retinal coordinates, yet our visual experience is stable across eye movements, implying a spatiotopic representation of information coding. Such coding relies on retinotopic coding plus knowledge of changes in the self's viewpoint/position due to eye, head or body movements. Early visual regions are retinotopically organised (Gardner, Merriam, Movshon & Heeger, 2008; Golomb, Mguyen-Phuc, Mazer, McCarthy & Chun, 2010; Crespi *et al,* 2011), whereas later visual regions show signs of spatiotopic coding (Zipser and Andersen, 1988). Neurons in parietal areas, for instance, show predictive coding of stimuli that will fall onto the receptive field once an eye movement has been made (Duhamel, Colby & Goldberg, 1992). Additionally, some properties of the ventral processing stream (e.g. large receptive field sizes, MacEvoy & Epstein, 2007; and position invariance, Grill-Spector & Malach, 2001) indicate spatiotopic coding.

Methods of psychophysics offer a way to examine what information is remapped across a saccade onto a spatiotopic reference frame through investigating the transference of aftereffects following perceptual adaptation. There is some evidence from such methods that spatially detailed stimuli are encoded in spatiotopic coordinates, and the degree of this encoding correlates with stimulus complexity, suggesting an increasingly spatiotopic representation along the processing hierarchy of the visual system (Melcher, 2005; Melcher & Colby, 2008), which is supported by some neuroscientific studies (Merriam, Genovese & Colby, 2007; Nakamura & Colby, 2002; although see Golomb & Kanwisher, 2011).

Determining whether specific adaptation-induced aftereffects are retinotopic or spatioptopic, therefore, offers a way to explore how specific visual attributes are encoded with relation to the processing hierarchy of the visual system. The motion aftereffect, for instance, is known to be fixed in retinotopic coordinates (Knapen, Rolfs & Cavanagh, 2009; Turi & Burr, 2012), reflecting its effect in early visual areas (possibly V1), whereas the *positional* motion aftereffect (the apparent change in spatial position following motion adaptation) is spatiotopic (Turi & Burr, 2012). It is predicted in the present study that, although aftereffects to mean orientation (i.e. the tilt aftereffect) are retinotopic, adaptation to orientation variance may reveal spatiotopic encoding. Although local orientation adaptation results in a retinotopic-specific tilt aftereffect (due to its operating in V1; Knapen, Rolfs, Wexler & Cavanagh, 2010), the encoding of orientation variance is likely to require the large receptive field properties of later neurons, which are more likely to encode spatiotopically (Melcher, 2005; Melcher & Colby, 2008).

One clear advantage of spatiotopic encoding is to align the activity within visual maps in a common reference frame that can be compared across modalities. The statistical computations that are required for orientation variance estimation are likely to be important for object recognition, specifically in determining the spatial regularity of surface features; information which can then be supplemented, for instance, through haptic input. If information regarding orientation variance is indeed encoded at a spatiotopic level of processing in visual cortex, then this information should be remapped around the time of a saccade to remain aligned with a location in space. Contrarily, if early retinotopic areas are sufficient for the coding of orientation variance, then any perceptual aftereffects should be restricted to the retinal coordinates in which they were induced. In order to dissociate these reference frames, it is necessary to invoke an eye movement between the presentation of the adaptor and test stimuli. In experiment three, therefore, observers fixated either above or below the adapting stimuli and, when prompted, moved their gaze diagonally either to the left or right of centre, before either returning to the original fixation position or to the opposite fixation location (above/below). Thus, the test stimuli were presented to one of four locations relative to the position of the adapting stimuli:

1. Same retinal, same spatial coordinates ("Full" condition)
2. Same retinal, different spatial coordinates ("Retinotopic" condition)
3. Different retinal, same spatial coordinates ("Spatiotopic" condition)
4. Different retinal, different spatial coordinates ("None" condition)

Finally, as attention is important for spatiotopic encoding of information (Crespi *et al,* 2011), subjects were encouraged to attend to the textures during the adaptation sections by carrying out a secondary task of indicating when one of the textures was presented with a relatively lower luminance contrast.

4.7. Experiment 3 – the reference frame of orientation variance adaptation

4.7.1. Methods

Participants

4 observers took part in the study, including the author. Participants gave their written, informed consent.

Equipment and stimuli

See experiment 1. In this experiment, gaze fixation was recorded using an infrared video eye-tracker (Cambridge Research Systems) with a sampling frequency of 250 Hz.

During adaptation, the pair of textures was always presented in the centre of the screen and, in the test phase, was presented either in the centre of the screen or in the upper or lower half (depending on the testing condition).

Experimental conditions and design

Each subject completed 4 testing sessions, factoring the side (left/right) and variance (low/high) of the adapting texture. Testing was conducted over a period of four days, with a maximum of one session per day. The order of the four sessions was counterbalanced across subjects, and half of the subjects fixated above the adapting stimuli, and half below. Each session consisted of four interleaved conditions, in which the test stimuli were presented at different locations relative to the adapting stimuli. Thresholds acquired for these conditions were averaged across the left and right adaptation side conditions. Thus, individual threshold estimates were gathered for four conditions of test stimuli location, individually for adaptation to low and high levels of variance.

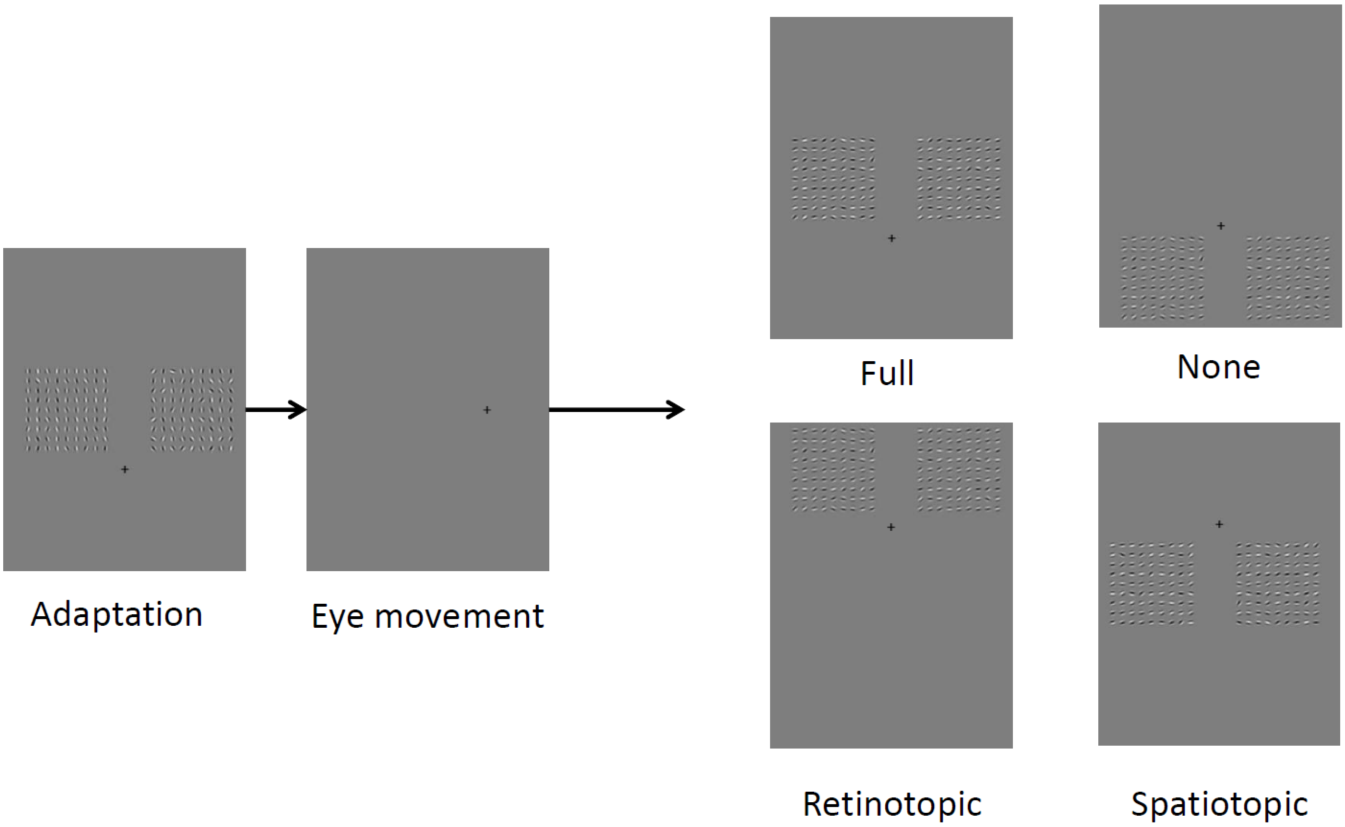

Figure 4.3 Illustration of the conditions from experiment 3. During adaptation, half of the subjects fixated above and half below (shown here) the adapting stimuli. The figure shows adaptation to high variance on the right, and medium on the left. Following a period of adaptation, an eye movement was required either to the left or right (shown here) of centre. Four stimulus conditions were interleaved in each block; in two of these conditions fixation returned to the original position for test, and in the remaining two fixation moved to the opposite side of the adapting stimuli (either above or below) for test. The test stimuli then appeared either above or below the new fixation point. Thus, the combination of the fixation position and test stimuli location determined the reference frame (labelled in the figure).

During adaptation, the fixation cross was presented 1.8° either below the adapting stimuli (measured from top/bottom of textures) for half of the subjects or above for the remaining half. The test stimuli were presented after a period of adaptation. Four experimental conditions were included (see **figure 4.3**), in which the retinotopic and spatiotopic position of the test stimuli relative to the adapting stimuli were independently varied. Dissociating these conditions required subjects either to relocate their fixation to the opposite side of the adapting stimuli or to maintain fixation at the same location, and then to present the test stimuli either above or below the new fixation point. In order to equate for the number of eye movements across the four conditions, however, two eye movements were required to be made in each condition regardless of whether fixation changed for the testing period or not; following adaptation, fixation moved diagonally either to the left or right of centre (equally likely) before relocating to the test position. The "full" aftereffect was measured when subjects relocated their fixation to its original position and the test stimuli were presented in the same location as the adapting stimuli. The "retinotopic" aftereffect was measured when subjects changed fixation to the opposite location and the test stimuli were presented on the opposite side of fixation to the adapting stimuli. The "spatiotopic" aftereffect was measured when subjects changed fixation to the opposite location and the test stimuli were presented in the same location on the screen as the adapting stimuli. The final control condition ("none") was measured when subjects relocated their fixation to its original position and the test stimuli were presented on the opposite side of fixation to the adapting stimuli. The purpose of this control condition was to correct for any retinotopic spreading of the adaptation effect that may explain any effect found in the spatiotopic condition that was not due to spatiotopic coding.

Trial sequence

Subjects fixated a cross either below or above the adaptation textures, and pressed a button to begin a session. Each session began with a 60s adaptation procedure consisting of 30 stimulus flashes, each lasting 1s and separated by an interval of 1s. Each flash contained the medium variance texture and the "adapting texture". For each adaptation flash, the textures were constructed *de novo*, with newly determined random mean orientations and each orientation drawn again from their relative distribution. Following adaptation, subjects were instructed by a warning tone to anticipate an eye movement. Following this (1s), the fixation cross would move diagonally to 5.7° either to the left or right of the centre of the monitor. The duration of this new fixation lasted 500ms, following which the cross either returned to its original position above/below the location of the adapting stimuli or moved to the opposite position below/above the stimuli, depending on the testing condition. Following an additional period of 750ms and a second warning tone, the test stimuli flashed onscreen for 500ms either above or below the new fixation point, again depending on the testing condition. Subjects indicated which of the two textures appeared to contain the most orientation variance by pressing one of two keys, following which the fixation cross would relocate to its original position (if necessary), marking the end of one trial. Between trials, subjects were presented with a further top-up adaptation consisting of 3 flashes, and after every 25th trial, a prolonged top-up adaptation of 30 flashes.

Trials were presented using eight randomly interleaved one-up one-down staircases; each of the four conditions was represented by two staircases per session. The test stimulus texture corresponding to the side of the adapting texture remained constant at the medium variance level, and each staircase automatically adjusted the variance of the contralateral texture until it reached an estimate of the subject's perceived subjective equality (PSE) between the two textures. In one staircase per condition, this texture began the session with a higher variance to the texture on the adapting side, and in the second staircase began with a lower variance (± 5 standard deviation units). The variance was incrementally increased/decreased initially by a magnitude of 2 standard deviation units (first 4 trials per staircase) and from thereon 1 standard deviation unit, until a criterion of 6 reversals had been met. Threshold estimation was acquired by averaging the final 4 reversal points.

As attention is important for spatiotopic encoding of information (Crespi *et al,* 2011), subjects were encouraged to attend to the textures during the adaptation sections by carrying out a secondary task. This was to indicate, by pressing a button, if one of the textures was constructed of Gabor patches of a slightly lower contrast than normal (60% Michelson contrast). This occurred with a probability of 0.1 on each adaptation flash, and was equally likely to be the left or right texture.

Additionally, eyetracking was carried out to ensure that subjects tracked the fixation cross during the intermittent relocation period following adaptation; if the eyetracker did not detect that the subject had fixated following 200ms or that fixation subsequent to this was not within a 1° radius of the cross then data from that trial was disregarded, and the experiment continued without collecting data from that trial. Additionally, if fixation was not recorded to be within a radius of 1° of the cross during the time in which the test stimuli were onscreen, the trial was also disregarded.

4.7.2. Results

The online fixation monitoring ensured that trials were not counted if subjects moved their gaze away from the fixation cross. To ensure that subjects were nonetheless covertly attending to the adapting stimuli, their hit rates and false alarm rates in detecting the contrast decrement which occurred, on average, once every ten flashes were compared. Hit rates (as a proportion) for each subject were 0.97, 0.97, 0.77 and 0.97, whereas no subject exceeded a false alarm rate of 0.01. From this, it is concluded that subjects were sufficiently attending to the adapting stimuli.

Thresholds collected from sessions in which the adapting texture was present on the left or on the right were averaged to give a single threshold estimate for each testing condition, and this was done independently for adaptation to low and high variance. Thus, each threshold estimate is taken from an average of four staircases, and each threshold value represents the shift in variance (in σ units) that is required to obtain PSE. These data are shown in **figure 4.4**. A 2 x 2 repeated measures ANOVA was run separately for the high- and low-variance adaptation conditions, with the relative position of the adapting and test stimuli in retinotopic coordinates (same vs. different) and spatiotopic (same vs. different) as factors (“Full”: same, same; “retinotopic”: same, different; “spatiotopic”: different, same; “none”: different, different). For low-variance adaptation, a significant main effect of retinotopic coordinates was found ($F_{(1,3)}$=16.43, p=0.027), with no main effect of spatiotopic coordinates ($F_{(1,3)}$=0.74, p=0.45) and no interaction ($F_{(1,3)}$=0.32, p=0.610). The same result was found for high-variance adaptation, with a significant main effect of retinotopic ($F_{(1,3)}$=16.12, p=0.028) but not spatiotopic ($F_{(1,3)}$=2.07, p=0.246) coordinates, and no interaction ($F_{(1,3)}$=0.33, p=0.608). Importantly, through examining **figure 4.4** it is confirmed that following adaptation to either low or high variance, respectively, the perceived variance of the test texture is either increased or decreased (as shown by the positive and negative aftereffects shown in the figure).

Additionally, it is important to assess the role of retinotopic spreading, specifically in the spatiotopic condition, by subtracting the magnitude of the “none” PSE from all other PSEs. The result of this is shown in **figure 4.4**, and clearly there is very little residual effect in the spatiotopic condition, whereas in the remaining retinontopic conditions it is still substantial. These results suggest that adaptation to orientation variance occurs in a strictly retinotopic reference frame.

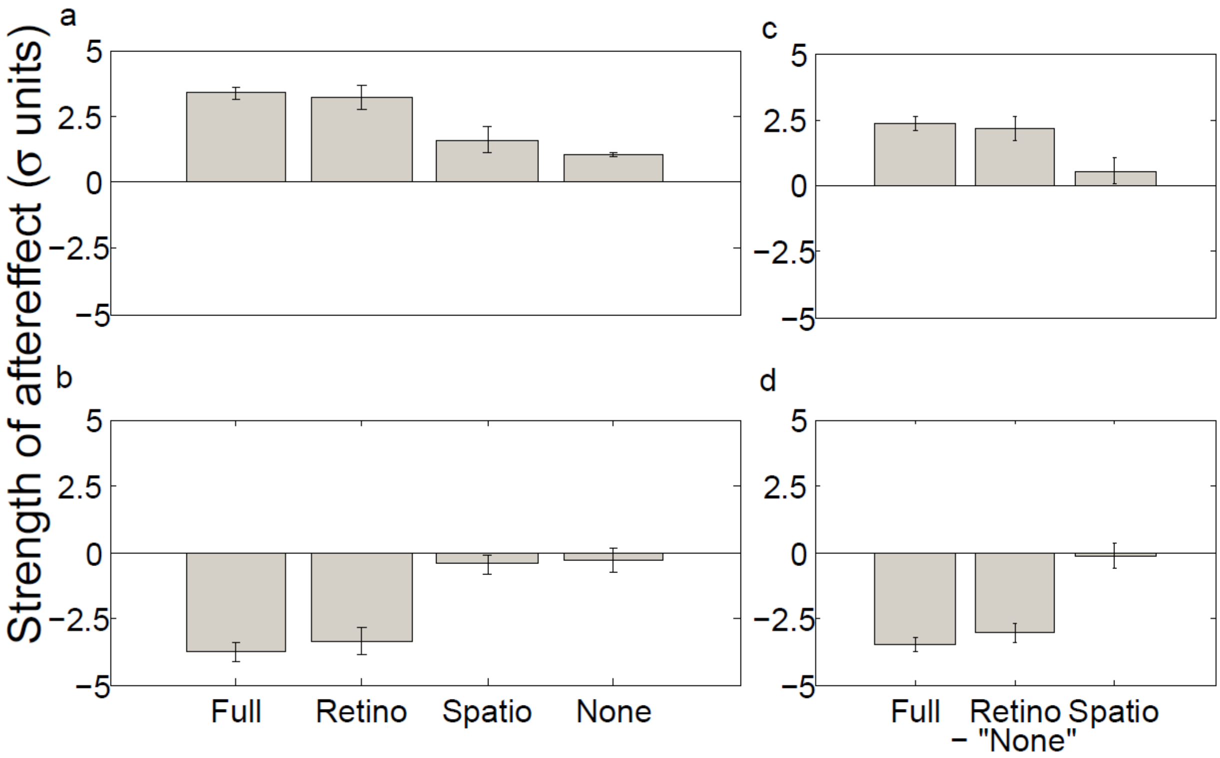

Figure 4.4. Results from experiment 3. Each data point represents the average threshold across four observers. Results from low adaptation are shown in **a**; those from high adaptation are shown in **b**. The strength of the aftereffect measurement is determined by the increment/decrement in orientation variance required to null the adaptation effect (PSE). In both cases, the full and retinotopic PSEs are clearly of the largest magnitude compared to the spatiotopic and nonspecific effects. When the "none" PSE magnitude is subtracted from the other PSEs to account for retinotopic spreading (**c** and **d**), the spatiotopic effect is marginal, whereas the other two conditions still show substantial aftereffects. Error bars show +/- 1 SEM, with between-subject variance removed.

4.8. Interim Discussion

In experiments one and two, selective adaptation aftereffects were used to reveal that orientation variance is an explicitly encoded dimension present in the visual system. Following adaptation to a texture of low variance, a subsequently viewed texture of a medium variance was perceived to be of a greater variance, whereas adaptation to high variance led to a reduced perception of variance in the medium texture. As shown in experiment three, however, this observed aftereffect was only specific to the two conditions in which the test stimuli appeared at the same retinal location as the adapting stimuli, and no evidence was found for the effect in spatiotopic coordinates. If this retinotopic specificity of the variance aftereffect indicates early cortical processing of orientation variance (Melcher, 2005; Melcher & Colby, 2008; Merriam *et al*, 2007; Nakamura & Colby, 2002), then patient MS, described earlier in this thesis, should retain at least some capacity to discriminate the level of variance between textures. MS is unable to make some region-based discriminations, including surface colour and texture of objects, which is most likely due to the bilateral damage to his medial occipitotemporal cortex, which has been shown in normal observers to correlate with the discrimination of such regional properties (Cavina-Pratesi, Kentridge, Heywood & Milner, 2010b). The early cortex in his left hemisphere, however, is spared, in which the retinotopic organisation of orientation-selective neurons may be sufficient to encode orientation variance. Experiment four tested whether MS retained sensitivity to orientation variance.

4.9. Experiment 4 – orientation variance discrimination in patient MS

4.9.1. Methods

Participants

Patient MS has bilateral damage to his ventromedial occipitotemporal cortex and is profoundly achromatopsic and prosopagnosic as well as having visual object agnosia. Damage to the left hemisphere includes the temporal pole, parahippocampal and fourth temporal gyri of the temporal lobe, the collateral sulcus, and the mesial occipitotemporal junction. His achromatopsia is most likely explained by the damage to the lingual gyrus and anterior collateral sulcus, typically associated with the human “colour area”. Damage in this vicinity is also likely to be the cause of his impairment in discriminating surface properties of objects more generally (Cavina-Pratesi *et al*, 2010b). His ability to perceive object form remains intact, which is likely the result of his spared lateral occipital complex (LOC; Cavina-Pratesi *et al,* 2010b). The damage in his right hemisphere is more extensive than that in the left, including primary visual cortex resulting in a homonymous hemianopia with macular sparing, whereas his occipital lobe in his left hemisphere is largely intact. For a more extensive case description of MS, see Heywood, Cowey and Newcombe (1994). Testing was conducted at the University of Durham’s Psychology Department. MS was 63 years of age at the time of testing.

Three non-age-matched control participants took part in the same experiment as MS.

Equipment and procedure

Equipment was the same as in the previous experiments. MS viewed the display monitor at a distance of roughly 80cm and was not required to fixate or to use a chin rest.

Stimuli consisted of three separate textures of 10 x 10 Gabor patches. Each Gabor patch measured 0.7° in diameter and had a spatial frequency of 1.4 cyc/degree. Each was separated from its neighbours by 0.3°. Thus, each texture measured 9.7° in width and in height. The phase of each Gabor was determined randomly from the full 360° cycle. The three textures were aligned vertically in the centre of the display, with a distance of 1.4° separating each of the top and bottom textures from the middle.

The orientations within each of the textures were drawn from a Gaussian distribution with a particular mean and variance. The respective mean values of the three textures were determined by first assigning one of the textures with a randomly determined value from the full 360° cycle. Of the two remaining textures, one was assigned this value +45° and the other +90°. Thus, each texture had a unique mean orientation with respect to the rest and this was novel on each trial. The baseline variance was chosen to be 7^2, which was assigned either to one or two of the three textures, with the remaining texture(s) being assigned a variance either of 12^2 (high difficulty), 20^2 (middle difficulty) or 30^2 (low difficulty). Thus, one of the textures was odd with respect to the other two, but it was not simply the case that this odd texture always had the largest (or smallest) variance; it was equally likely on each trial that the odd texture would have more or less variance relative to the other textures. MS was instructed to indicate which texture, either the top of bottom, appeared to be the odd one out. Importantly, with this method, MS could not complete the task on the basis of local orientation comparisons or by analysing only one of the textures; MS was required to make variance estimations of at least two of the textures. **Figure 4.5** provides an illustration of the stimuli.

The stimuli were presented onscreen until MS verbalised his response, at which point the experimenter pressed the appropriate response key. Five blocks were conducted, each containing 10 repetitions of each of the three discrimination levels, amounting to 30 trials per block.

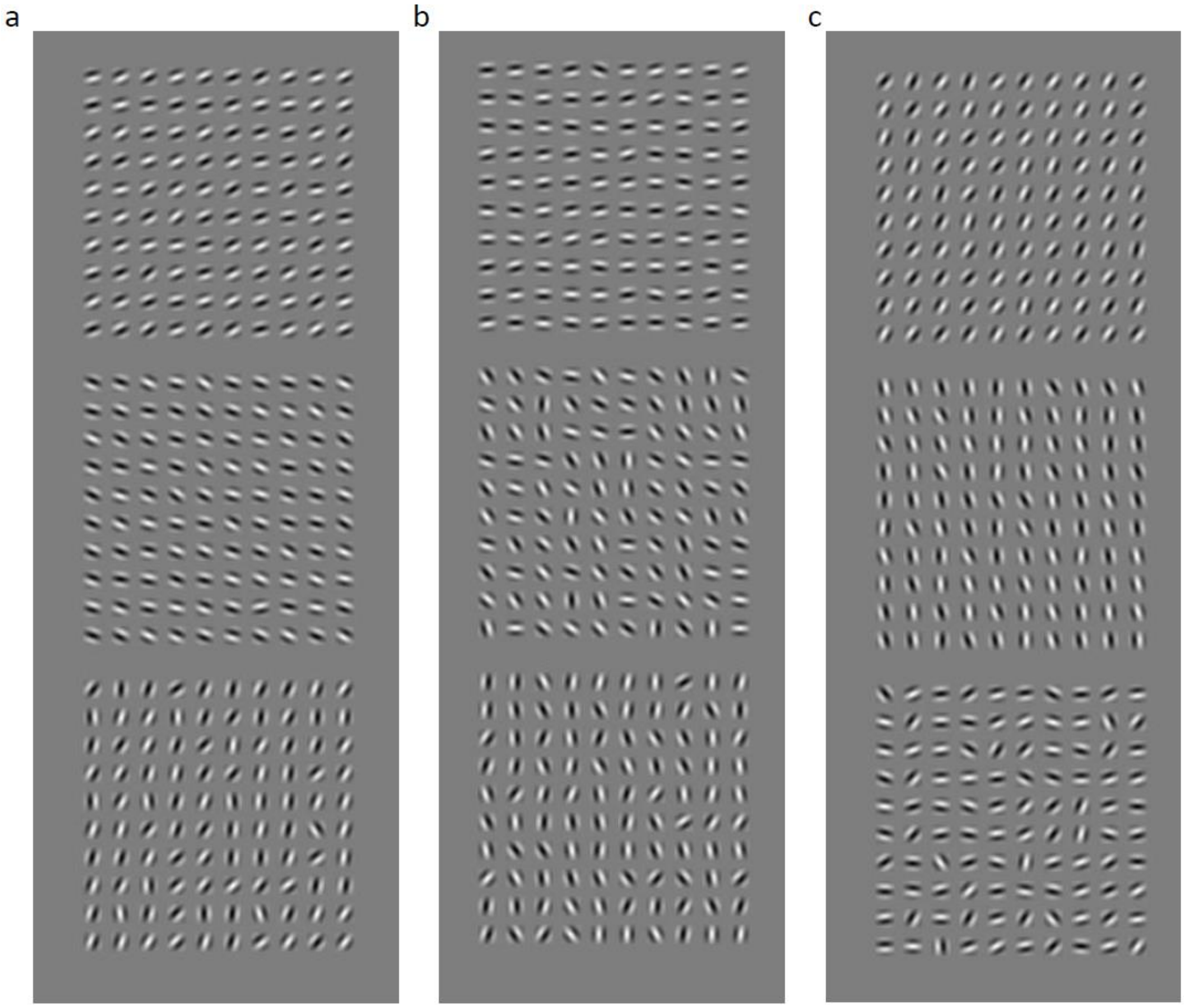

Figure 4.5. Example stimuli used in experiment 4. The task was to identify the odd texture (top or bottom) in terms of its variance (irregularity). The three difficulty levels are shown in **a-c**, with decreasing difficulty level left to right. Each texture had a unique mean orientation relative to the other two, which was randomly determined across trials, thus preventing scrutiny of the stimuli on a local scale.

4.9.2. Results

For the high, middle and low difficulty levels, respectively, MS produced the following hit rates: 30/50 (60%; p=0.101), 37/50 (74%; p<0.001) and 42/50 (84%; p<0.001), shown as percentage scores in **figure 4.6**. Normal observers (N=3) on the same task did not produce errorless performances and showed the same rising trend in performance with decreasing task difficulty (90%, 97.3% and 100%). These do suggest that, although MS was clearly impaired relative to normal non-age-matched controls, he shows some sensitivity to the property of orientation variance in the textures.

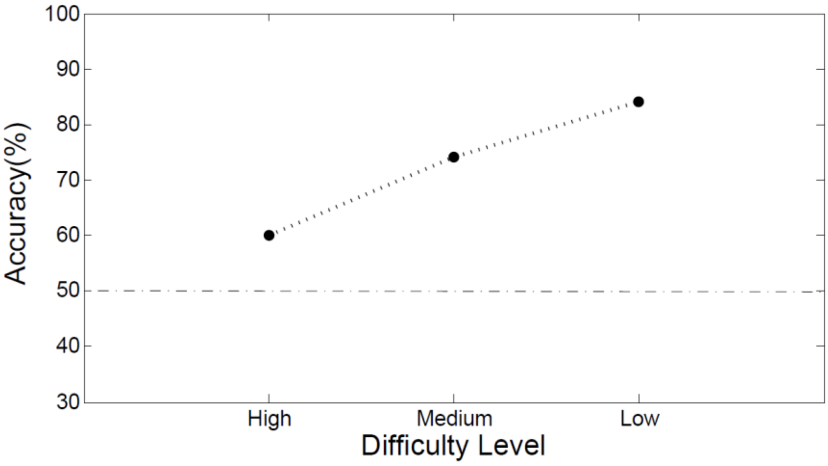

Figure 4.6 Results from experiment 4; MS's variance discrimination task. MS performed significantly greater than chance in the medium- and low-difficulty levels and with progressively more accuracy with increasing variance difference between the discriminanda. This suggests that he may retain an ability to perform the second moment regional estimations that allow the encoding of variance.

4.10. General discussion of study three

The effect of variance adaptation is clearly demonstrated in experiments one and two. The results suggest that what has been adapted is a specific mechanism that is selective to orientation variance, a second moment statistic of regional orientation information. Importantly, due to the random allocation of mean orientation throughout experiment one, the effect can not be explained on the basis of an undesired effect of adaptation to local orientation (i.e. the tilt aftereffect), which might otherwise have produced a change in in the stimuli that could have been misperceived as a change in variance. Additionally, because of the 90° disparity in mean orientation between the adapting and test stimuli in experiment two, the results imply that variance is encoded at a general level of processing, at which point the low-level orientation statistics do not matter. Thus, the adaptation effect is robust and is specific to the statistic in question (variance). This echoes results found following adaptation to other statistical properties (Durgin, 2001; Durgin & Huk, 1997; Motoyoshi *et al, 2007*; Corbett *et al,* 2012), and findings that regional information of a set, or texture, is encoded independently of individual local components (Ariely, 2001; Corbett *et al*, 2012).

Similar findings regarding image statistics have come from other lines of research. In the case of visual crowding, for example, in which objects in the periphery become difficult (or impossible) to individuate due to the presence of neighbouring flanks, it has been shown that, despite the inability to individuate single orientation elements, observers can nonetheless make judgments based on their statistics, specifically their mean and variance (Parkes, Lund, Angelucci, Solomon & Morgan, 2001; Solomon, 2010). Such findings, along with the work of Morgan and colleagues (2008), point to a mechanism within the visual system that explicitly encodes the variance of a set of oriented lines. Indeed, the results of experiments one and two suggest a model in which the neural representation of variance is determined at least by two channels, one tuned broadly for low variance and one for high variance, and that the perception of variance is determined by the complementary balance of activity within these channels. An alternative theory is, of course, that there are multiple narrowly tuned channels, much like those that determine sensitivity to the property of spatial frequency.

The adaptation aftereffect, however, as shown in experiment three, is limited to a retinotopic reference frame; aftereffects were not found for test textures that appeared in the same spatial location as the adaptors independently of retinal location. This is also true of the tilt aftereffect, which follows adaptation to a single orientation (Knapen *et al*, 2010). Knapen and colleages (2010) argue that it would be unlikely to find spatiotopic aftereffects of orientation adaptation, due to it being carried out in V1; remapping of the adaptation would require that the modified, adapted state of neurons be transmitted horizontally through the brain in any direction depending on the impending eye movement, necessitating a very dense connectivity in lower visual areas for which there seems to be no neurophysiological evidence. It was hypothesised in the present study, however, that aftereffects specific to variance adaptation may reveal spatiotopic encoding, due to the large receptive fields that may be required to achieve such large-scale estimations. These large-scale receptive fields are to be found downstream from primary visual cortex, in areas which display progressively more signs of spatiotopy (Merriam *et al*, 2007). In the present study, however, no evidence for a spatiotopic aftereffect was found. This may indicate that variance perception is achieved and represented at a much earlier level in the visual system than previously expected.

There is, however, some contention as to whether mid- to high-cortical areas indeed show an increased sensitivity to spatiotopic encoding, as claimed, for example, by Melcher (2005). Only very small, if any, spatiotopic aftereffects are often found for stimulus properties that are thought to reflect the processing of spatiotopically-tuned cortical areas (e.g. Wenderoth & Wiese, 2008). Some evidence from imaging also suggests that the vast majority of processing in low- and high-level areas is strictly retinotopic (Gardner *et al*, 2008). These findings, however, may only reflect differences in the stimuli used and the lack of directed spatial attention towards the stimuli. In Gardner *et al's* (2008) study, for instance, the stimuli were presented parafoveally whilst subjects performed an attentionally demanding task at fixation. It is now known that selective spatial attention towards a stimulus is a crucial factor in determining its spatiotopic encoding in mid- to high-level areas (Crespi *et al*, 2011). The failure to find spatiotopic aftereffects in the present study can not be explained by the claim that subjects were not attending to the adapting stimuli, as they performed very well in a secondary task that required them to make judgements of the Gabor patches' contrast[4].

The results from experiment four adds to the theory that orientation variance is encoded relatively early in the visual system, as patient MS is shown to be sensitive to the level of orientation variance in the textures. Although MS performed poorly relative to young control participants[5], he performed significantly above chance in the two least difficult conditions and showed increasing accuracy with decreasing discrimination difficulty. Furthermore, he could not have achieved this on the basis of local orientation comparisons between the textures, as each texture was assigned a mean orientation which was different to the other two and randomly allocated across trials, therefore requiring an estimation of variance independently of mean orientation for at least two of the three textures in order to complete the task. It may be, of course, that accurate variance discrimination requires a collection of cortical areas, some of which are not functional in MS. Alternatively, however, it may be his intact early cortex which may be sufficient. In light of recent evidence, this may not be entirely surprising. Joo, Boynton and Murray (2012), for instance, have recently

---

[4] A number of recent studies have also failed to demonstrate spatiotopic aftereffects despite including a secondary attention task (Afraz & Cavanagh, 2008; 2009; Knapen *et al*, 2009).

[5] His performance in the task is undoubtedly affected by his homonymous hemianopia, which restricts the number of orientation samples he can make at any one time.

documented how early visual cortex is sensitive to long-range contextual patterns of orientation. Specifically, if a vertical Gabor patch is flanked by two horizontal Gabor patches (i.e. H-V-H), then sensitivity (and neural activity) is greater for this central Gabor patch than if an additional vertical Gabor patch is placed either side of the flankers (i.e. V-H-V-H-V). The authors hypothesised that, because in the latter condition the central Gabor patch is part of a long-range pattern, and therefore does not deviate from its context, the visual system is less sensitive to it. Importantly, they showed event-related potentials that that strongly suggested this effect to be mediated in the earliest components of visual cortex. This is also somewhat consistent with other work involving patient HJA (Allen, Humphreys & Bridge, 2007), whose preserved striate cortex presumably allowed him to successfully average, albeit not perfectly, the orientations of a number of Gabor patches. The results from the present study suggest that even the second moment of orientation processing may be found in early retinotopically organised areas of visual cortex.

5. Study Four - The role of extrastriate cortical areas in representing cue-invariant 2nd-order segmentation

5.1. Abstract

The visual system's ability to represent object shape depends largely on its mechanisms to detect local discontinuities in visual features (e.g. luminance, colour, texture), a process known as segmentation. Computations of different types of local contrast, however, are undertaken by different populations of neurons and, over time, the type of cue that defines the form an object can often be subject to change. The temporal integration of information from different cue sources is thus essential in establishing a stable perception of an object in a changing environment. In this study, a novel method of flicker-defined-form was applied, wherein the cue contrast that defines the contours of a figure rapidly alternated between one type of 2nd-order cue (orientation) and another (spatial frequency). Experiment one showed that segmentation of a figure in this manner dissociated from the regional discrimination of the same stimuli (i.e. when the information at the locations of contrast was masked by black borders), and that the magnitude of this difference was no different than when the figure was defined by a contrast of a single constant cue, implying the successful rapid integration of 2nd-order visual features by the visual system. In order to ascertain whether this perceptual integration process necessitates additional cortical involvement relative to single-cue segmentation, experiment two compared sensitivity to the stimuli selectively in the upper and lower visual fields. Segmentation of a target defined by an alternating cue showed a large lower visual field advantage in comparison to single-cue segmentation. This asymmetry implies the recruitment of a cue-invariant visual integration area beyond the level of primary visual cortex. In a third experiment, patient DF, who has bilateral lateral occipital cortical damage, was shown not to be selectively impaired in segmenting by a rapidly alternating cue type. Taken together, these results suggest a cue-integration area for the purposes of segmentation that is likely to be found subsequent to primary visual cortex and prior to lateral occipital complex.

5.2. Introduction

In detecting the boundaries of objects in the environment, the visual system's primary adaptation lies in rapidly detecting discontinuities in visual information; a process often termed *segmentation*. 1st-order segmentation can refer to detecting boundaries on the basis of uniform luminance or colour. Other forms of segmentation, however, require 2nd-order comparisons (e.g. orientation and spatial frequency), yet these can nonetheless be perceptually very salient (Chubb, Olzak & Derrington, 2001; Motoyoshi & Nishida, 2001; Bergen & Adelson, 1988). The process of segmentation is automatic and results in the perception of boundaries around the segmenting regions, thus allowing figure-ground relationships to be determined and objects to be identified and acted upon. This contrasts somewhat with the ability to *discriminate* visual regions, which aids in other aspects of object recognition but is a slower and more effortful process. For present purposes, the word discrimination is reserved for instances in which two regions of information appear different but do not necessarily segment.

The visual system's ability to segment has only traditionally been studied, however, when the cue type which constitutes the contrast remains constant over time (single-cue segmentation). This is informative for understanding the low-level functional properties of edge localisation, most of which occur in striate cortex (Dai & Wang, 2011; Kentridge, Heywood & Weiskrantz, 2007; Knierim & van Essen, 1992; Sillito, Grieve, Jones, Cudeiro & Davis, 1995), but does not offer much insight into subsequent integration stages beyond the level of single-cue segmentation. Some researchers have investigated in great depth how multiple simultaneously available cues are integrated to achieve complex visual goals (e.g. depth perception and segmentation; Landy & Kojima, 2001; Saarela & Landy, 2012). What is less clear in our understanding, however, is how the visual system integrates separate cue types as information that defines an object changes over time. This process should not be overlooked in efforts to understand segmentation, as it is a common aspect of perception that may arise from sudden changes in environmental conditions or changes in an observer's perspective. Despite these factors, we achieve perceptual constancy and experience a stable world in which we are able to track objects and to recognise them in different contexts. Perception, therefore, must become cue-invariant at some stage. The present study concerns two types of 2nd-order segmentation cue (orientation and spatial frequency) and aims to address where in the hierarchy of visual processing information from these two different cue sources is rapidly integrated over time.

Flicker-defined form (FDF) was applied in the present experiments to isolate the segmentation component of perception. FDF is a phase-reversing masking technique first developed in a study by Rogers-Ramachandran and Ramachandran (1998), which demonstrated that when the luminance of two abutting light and dark regions is continually reversed above a particular frequency (~ 8 Hz), it becomes no longer possible to distinguish the dark region from the light region at any one time, yet a contour is still clearly visible between the two[6]. This technique exploits the fact that segmentation persists at a higher frequency than discrimination and, importantly for this study, is also applicable to 2nd-order stimuli (Norman, Heywood & Kentridge, 2011). The notion that there exist separate mechanisms for segmentation and discrimination of the same stimuli is further supported by experiments employing various other techniques of visual masking (e.g. Breitmeyer, Kafaligönül, Ögmen, Mardon, Todd and Ziegler, 2006; Motoyoshi, 1999).

The experiments reported here used a novel variant of FDF in which the type of cue that defines the contrast rapidly alternates between one type of 2nd-order cue (orientation) and another (spatial frequency), whilst the locations of contrast remain constant. If the visual system is able to segment a figure based on such a contour defined by an alternating-cue, and can achieve this at a higher temporal frequency than the ability to discriminate the regions (i.e. as in typical FDF), it would suggest the involvement of a visual integration stage beyond the level of simple single-cue segmentation that must be cue-invariant, i.e. in which there is neural synergy between the encoding of different visual features.

In the following experiments, subjects completed a localisation task in which a homogenous background of Gabor patches contained a smaller target region defined by some feature contrast. This feature contrast would either remain as one type (single-cue

---

[6] The technique of flicker-defined form is proposed by some to isolate the magnocellular subdivision of visual processing (Rogers-Ramachandran & Ramachandran, 1998), although this is contested by others (Skottun and Skoyles, 2006). The present study does not attempt to make this claim for the stimuli used in these experiments; the technique is only used as a method to dissociate the processes of segmentation from those of discrimination, regardless of their relative contribution from the magnocellular subdivision of visual processing.

condition; orientation contrast or spatial frequency contrast) or rapidly alternate between the two (alternating-cue condition) on a continuous cycle.

Experiment one was designed to demonstrate that the individual cues (orientation and spatial frequency) can be successfully integrated to rapidly segment the figure. The rationale was that if the cues are successfully combined by a cue-invariant integration area, then there should be no decrease in the ability to segment a target defined by a rapidly alternating cue relative to one defined by a single cue. To determine this, flicker frequency thresholds were measured for segmenting a target defined by either a single cue or an alternating cue, along with the same thresholds for discrimination of the target regions (i.e. when segmentation is prevented by the presence of coincident border masks).

5.3. Experiment 1 –cue-invariant segmentation in flicker-defined-form

5.3.1. Methods

Participants

10 participants took part in experiment one. Participants were either undergraduate or postgraduate students at Durham University's psychology department. They were compensated for their time either financially or with participant pool credits. All had normal or corrected-to-normal vision; none were experienced psychophysical observers.

Stimuli and procedure

Stimuli were generated using a Cambridge Research Systems VSG 2/5 Graphics System and were presented on a gamma-corrected ViewSonic 17'' display monitor viewed at a distance of 41cm (subjects rested their head on a chin rest). The background had a luminance of 50cdm^{-2}. The screen resolution was set to 1024 x 768 pixels with a refresh rate of 100 Hz.

Subjects fixated a central cross. Following a warning tone, a lattice (20.3° in width and 10.0° in height) centred on the fixation cross consisting of 20 x 10 uniformly positioned Gabor patches appeared. The central 2 x 2 Gabor placeholders were empty, to allow subjects to fixate the central cross with minimal distraction from neighbouring stimuli. Each Gabor had a diameter of 0.70° and was separated by its neighbours by 0.33°. The Michelson contrast of each Gabor was set at 90%.

The temporal sequence for each trial began with a mask (lasting 30 ms), followed by a continuous stimulus cycle of 6 frames (lasting 1000 ms; depicted in **figure 5.1**), and followed by a final mask (30 ms). Mask frames and stimulus frames were comprised of Gabors in the same positions, but the Gabors in the mask would each comprise a randomly and individually determined orientation from the uniform distribution within the bounds of -12.5° and +12.5° about vertical, and a similarly determined spatial frequency within the bounds of 1.4 and 2.9 cycles/degree. Each Gabor's phase was randomly determined from a uniform distribution within the full 360° range (this was also true for Gabor patches present in the stimulus frames).

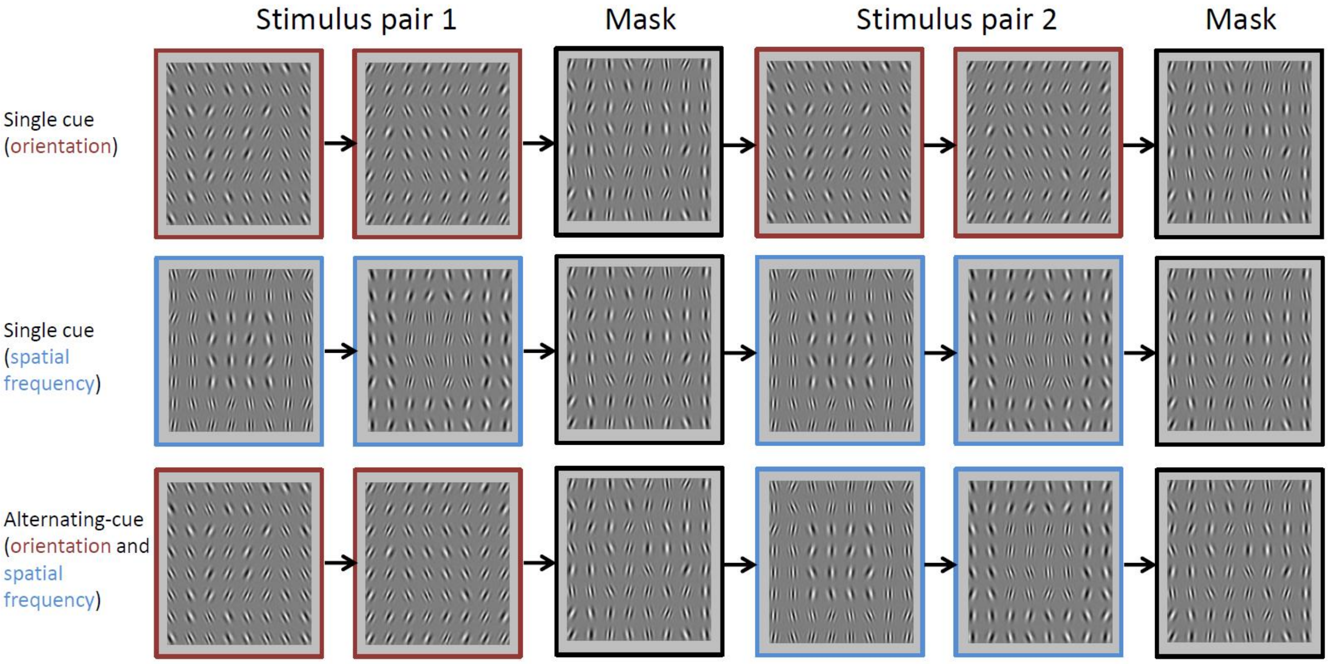

Figure 5.1 6-frame stimulus sequence (from left to right) for three conditions of contrast type (by row). In the top two rows, the target is defined by the same type of contrast in pairs 1 and 2 (either orientation or spatial frequency). In the third row, the target is defined by a different type of contrast in each pair; therefore if stable segmentation is to be achieved in this condition, the information from the different cue types must be temporally integrated. The mask frames were included to minimise interference between the stimulus pairs and to equate the changes across different conditions (see text for further details). Note that each box in the figure is a magnified section of only part of the display (the target), which would occur unpredictably in one of two locations on each trial.

The 6-frame cycle consisted of, in sequence, 1 set of 2 target frames (pair 1), 1 mask frame, 1 set of 2 target frames (pair 2) and a final mask frame. The target was a square region of the lattice measuring 4 Gabor spaces in width and in height which, for a given trial, was randomly chosen to be located on the left or on the right (measured 3 Gabor spaces from the edge of the lattice). The target, when present, was defined either by an orientation contrast or a spatial frequency contrast to the background and, depending on the condition, could be defined either by: a) an orientation contrast in pair 1 and an orientation contrast in pair 2; b) a spatial frequency contrast in pair 1 and a spatial frequency contrast in pair 2; or c) an orientation contrast in pair 1 and a spatial frequency contrast in pair 2. Thus, conditions a) and b) present a target that is defined by the same contrast type in both frame pairs, whereas condition c) presents a target that is defined by a changing contrast type (“alternating cue” condition). The order in which the two stimulus pairs were presented in condition c) was randomly determined for each trial.

Within each pair of target frames, both frames contained the same degree of contrast; however, across the two frames the phase of this contrast was reversed. In the condition of an orientation contrast, the background and target Gabor patches would assume orientations of -12.5° and +12.5° (a contrast of 25°) about vertical before reversing to 12.5° and -12.5°, respectively, with spatial frequency being determined equally in both cases by drawing each Gabor patch’s spatial frequency from a uniform distribution within the bounds of 1.4 and 2.9 cycles/degree (equivalent to the spatial frequency content of the mask). Similarly, in the condition of a spatial frequency contrast, the background and target Gabor patches would assume spatial frequencies of 1.4 and 2.9 cycles/degree (a contrast of 1.5 cycles/degree), before reversing to 2.9 and 1.4 cycles/degree, with orientation being determined equally in both cases by drawing each Gabor patch’s orientation from a uniform distribution within the bounds of -12.5° and +12.5° (again, equivalent to the orientation content of the mask). Following the 6 frame cycle a final mask appeared for 30 ms. See figure 1 for an illustration.

The phase-reversing procedure used in the two pairs of stimulus frames ensured that the target and background regions contained on average the same information when temporally fused, such that when the stimulus cycle was presented at a sufficiently high frequency, the ability to discriminate the regions fails with no substantial effect on the perception of the contour between the two (i.e. FDF). The intermittent masks between the two pairs of target frames were included to minimise stimulus interference between the two target frame pairs equally across the 3 conditions. For instance, without the mask frames, the change in orientation content in the single-cue orientation condition between the last frame in pair 1 and the first frame in pair 2 would be 25°. In the alternating-cue condition, this would only be 12.5°. The larger change in the former condition may have produced a larger masking effect, thus affecting subjects' detection thresholds, if the two conditions were not equated with the intermittent masks.

Following the offset of the stimuli, subjects indicated by pressing one of two buttons on a Cedrus button box whether the target appeared to the left or the right of fixation. An inter-trial interval of 1000 ms was included. In addition to completing the task as previously described, subjects also completed one condition in which two border masks were present for the duration of the trial. These consisted of two black boxes made of lines 1 pixel in thickness that each surrounded the two possible target locations. These ensured that when locating the target, subjects relied on discriminating the regional information within each area, rather than the contour formed by the contrast. Thus, the experiment was a 3 (contrast type) x 2 (border mask presence) within-subject design. All three contrast type conditions were blocked in a single session and subjects completed two sessions; one of the two sessions included the border masks and the order of these two sessions was counterbalanced across subjects. These two sessions were each preceded by a full practice session. Performance was monitored using an interleaved staircase procedure, and the frequency rate (measured in Hz) of the cycle[7] was determined by the subject's performance. Three 2-up, 1-down interleaved staircases were ran in a single session, with each staircase corresponding to one of the 3 contrast types (orientation, spatial frequency and alternating-cue). Frequency initiated at 5.6 Hz (9 frames per stimulus) and increased or decreased depending on performance by one frame. Staircase selection for each trial was randomly determined with equal probability, and each terminated after 8 reversals in performance. Threshold estimation was taken by averaging the final 4 points of reversal.

---

[7] The temporal frequency of the cycle does not refer to the *entire* 6-frame cycle; just the frequency of two of the frames in the cycle (i.e. one stimulus frame pair). The frequency of the entire cycle, therefore, is equal to this frequency divided by 3.

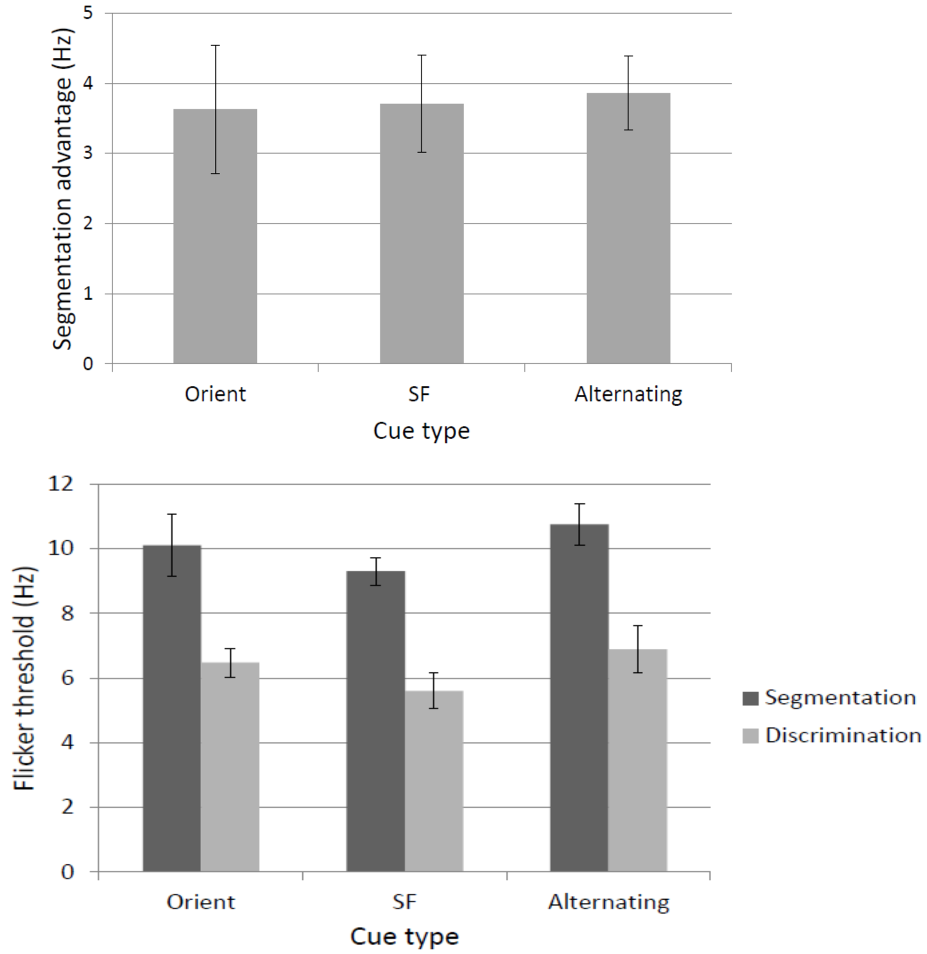

Figure 5.2 Mean segmentation advantage levels (upper panel) and raw threshold levels (lower panel) from experiment one. The segmentation advantage values are calculated by subtracting the subjects' flicker frequency thresholds in the discrimination condition (border masks present) from their thresholds in the segmentation condition (border masks absent). There is no drop in this magnitude for the alternating-cue condition relative to the single-cue conditions. Error bars show +/- 1 SEM, with between-subject variance removed.

5.3.2. Results

The magnitude of the "segmentation advantage" was calculated for each of the cue types by subtracting the discrimination thresholds (border masks present) from the segmentation thresholds (border masks absent)[8], giving a measure of how much the figure

---

[8] The approach of calculating the "segmentation advantage" was used because, in pilot runs of the experiment, it was found that subjects performed consistently better in the discrimination condition of the alternating-cue stimuli relative to the single-cue stimuli (which were matched for salience). The segmentation advantage thus represents how sensitive the observers are to the target regardless of their sensitivity in the relative discrimination conditions. The advantage found in the alternating-cue discrimination condition relative to that of the single-cue may reflect a change in the observers' decision criteria due to information from different channels being available to inform the discrimination, or it may be the effect of a dedicated cue-integration system that is more

segments. These values are shown in **figure 5.2** and were entered into a single-factor repeated measures ANOVA with three levels (orientation, spatial frequency and alternating cue type). There was no difference between the segmentation advantage levels of the three cue types ($F_{(2,\ 18)}$=0.027, p=0.973). This implies that when performing 2nd-order segmentation, the perception of the contour is represented at a stage at which there is synergy between the processing of orientation- and spatial frequency-defined contrast.

5.4. Interim discussion

The results from experiment one indicate that the visual system is capable of extracting contours that are formed by rapidly alternating cue types. This may indicate a synergistic stage in visual perception, in which information from different cue sources is integrated over time in order to subserve cue-invariant form perception. It is predicted that this cue-invariance would be found at a later stage in the visual cortex than that of single-cue segmentation, which is likely to occur as early as primary visual cortex (V1; Hallum, Landy & Heeger, 2011), and that a selective lower visual field (LVF) advantage would reveal this.

A LVF advantage has been found in a number of perceptual and behavioural experiments (e.g. Riopelle & Bevan, 1952; Rubin, Nakayama & Shapley, 1996; Levine & McAnany, 2005) and is believed to reflect differential processing of visual information in areas V2 and beyond (Horton & Hoyt, 1991; He, Cavanagh & Intrilligator, 1996). Relevant to the present study, 2nd-order segmentation does show a LVF advantage, although this is only very slight (Graham, Rico, Offen & Scott, 1999). Contrarily, segmentation by a rapidly alternating cue type is likely to involve higher cortical regions, as the earliest level in the visual system at which evidence of cue-invariance is found is in V2 (Leventhal, Wang, Schmolesky & Zhou, 1998[9]; Marcar, Raiguel, Xiao & Orban, 2000), and this should therefore

---

sensitive when information from multiple cue sources is available (Rivest & Cavanagh, 1996; Self & Zeki, 2005). Nonetheless, it is more important for the present study to show that there is a segmentation advantage relative to discrimination (as this is a measure of FDF), and that this advantage is equivalent in the three conditions of cue type.

[9] Leventhal *et al* (1998) did find that a very small proportion of V1 neurons do show cue-invariance, in that their activity correlates with the salience of a boundary irrespective of the cue that defines it.

be revealed through a LVF perceptual advantage. In experiment two, therefore, the same conditions were used as those in experiment one, but they were tested selectively in the UVF and LVF. The border condition remained such that the larger LVF advantage for the alternating-cue condition could be shown to be specific for segmentation (borders absent). This is important as it would suggest that the cue-invariant nature of form perception is specific to the processes that are solely related to *edges* and not *regions.*

5.5. Experiment 2 –lower visual field advantage for cue-invariant segmentation

5.5.1. Methods

Participants

10 observers participated in experiment two, including the author and one who previously took part in experiment one.

Stimuli and procedure

The stimuli used were the same as those in experiment one. Additionally, however, the Gabor lattice now included the 4 central Gabor patches that were missing in experiment one, as subjects did not fixate the centre of the lattice. Instead, the whole lattice could either appear 2.1° below or above fixation, with a high tone before stimulus onset indicating “above” and a low tone indicating “below”. This experiment was a 2 (border mask) x 2 (visual field) x 3 (contrast type) within-subject design. All 6 combinations of contrast type and visual field were blocked in a single session, and subjects completed 3 repetitions of a session without the border masks and 3 repetitions of a session with the border masks. The order of these was counterbalanced across subjects. All 6 sessions contained six 2-up, 1-down interleaved staircases, with each staircase corresponding to the 6 combinations of contrast type and visual field. Each staircase began at 4.5 Hz (11 frames per stimulus) and proceeded in the same way as in experiment one. Threshold estimation for a session was taken by averaging the final 4 points of reversal, and a subject’s overall threshold was taken by averaging the thresholds from the 3 repetitions.

5.5.2. Results

Frequency thresholds were entered into a 2 x 2 x 3 repeated measures ANOVA, with the factors border mask (present or absent), visual field (upper or lower) and contrast type (orientation, spatial frequency or alternating-cue). The main effect of border mask was significant ($F_{(1,9)}$=19.97, p=0.002), with subjects being able to detect the target at an overall higher frequency rate without the border masks (9.0 Hz) than with the border masks (6.6 Hz). Also, the main effect of visual field was significant ($F_{(1,9)}$=11.55, p=0.008), with subjects being able to detect the target at an overall higher frequency rate in the LVF (8.7 Hz) than in the UVF (6.9 Hz). The main effect of contrast type was not significant ($F_{(1.24,11.19)}$=1.21, p=0.308; Greenhouse-Geisser corrected). There was no interaction of border and visual field ($F_{(1,9)}$=0.18, p=0.685), border and contrast ($F_{(2,18)}$=3.25, p=0.063), or visual field and contrast ($F_{(2,18)}$=2.25, p=0.135). There was, however, a significant interaction of all 3 conditions ($F_{(2,18)}$=7.87, p=0.004).

To explore this further, the 2 conditions of border mask were analysed separately with 2 (visual field) x 3 (contrast) repeated measures ANOVAs. For the condition with the border masks present, there was only a significant effect of visual field ($F_{(1,9)}$=8.83, p=0.016). For the condition with the border masks absent, however, there was a significant effect of visual field ($F_{(1,9)}$=8.45, p=0.017) and an interaction between visual field and contrast (Greenhouse-Geisser corrected; $F_{(1.26,11.35)}$=5.95, p=0.027). Three planned paired t-tests were used to explore the magnitude of the visual field asymmetry for the 3 contrast types: whereas the average threshold was not significantly higher in the LVF relative to the UVF for spatial frequency contrast ($t_{(9)}$=1.38, p=0.200), it was for orientation contrast ($t_{(9)}$=2.44, p=0.037) and in the alternating-cue condition ($t_{(9)}$=3.97, p=0.003), with the effect being greater in the latter. **Figure 5.3** illustrates this: the LVF superiority effect is noticeably greater for alternating-cue segmentation compared to either type of single-cue segmentation. Importantly, this effect is absent in the condition with the border masks, indicating that this is an effect specific to segmentation, and not regional discrimination, of object shape.

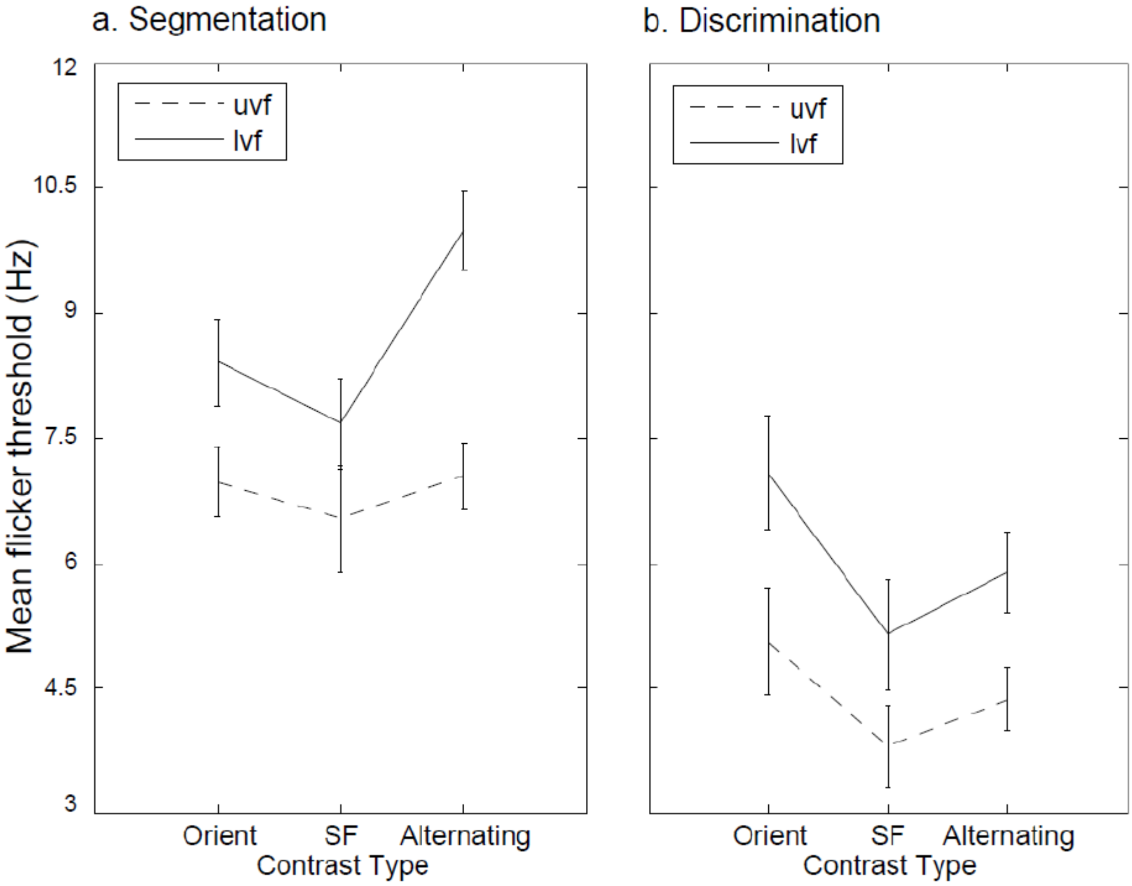

Figure 5.3 Mean flicker frequency threshold results from experiment two. When the target segments from the background (a) there is a noticeably larger LVF superiority in the alternating-cue condition compared with either of the single-cue conditions alone. Contrastingly, no such pattern exists for when the target region must be discriminated rather than segmented (b), as a similar LVF superiority is found in all three cases. Error bars show +/- 1 SEM, with between-subject variance removed.

5.6. Interim discussion

In the previous experiment, a general LVF advantage was found for all conditions. There was, however, a much larger advantage specific to the segmentation condition when the target was defined by an alternating-cue, potentially indicating the selective involvement of a cue-invariant integration stage at the level of V2 or beyond. A likely neural candidate for this role is the lateral occipital complex (LOC), a subsection of the ventral processing stream known to be highly selective for form perception. Imaging studies (Malach et al, 1995) and lesion cases (Goodale, Milner, Jakobson & Carey, 1991) together provide a strong case for this region being responsive to, and necessary for, the perception of object form, as well as being a likely cortical locus of object-based attention (de-Wit, Kentridge & Milner, 2009a). Most importantly, however, the neural activity observed in LOC is unvarying in response to changes in the type of cue that constitute the object (e.g. motion, texture; Grill-Spector, Kushnir, Edelman, Itzchak & Malach, 1998), is responsive to information from multiple sensory modalities (Kim & Zatorre, 2011) and shows greater activity in response to objects defined simultaneously by two cues as opposed to one (Self & Zeki, 2005), suggesting that it is highly cue-invariant. Additionally, LOC shows a LVF bias relative to the UVF (Sayres & Grill-Spector, 2008), which is likely to explain prominent LVF biases in form perception by illusory contours (Rubin et al, 1996), a process known to depend on LOC (de-Wit, Kentridge & Milner, 2009b). The activity within LOC in response to region-based information is slight in comparison to that which is edge-based (Shpaner, Murray & Foxe, 2009), which further suggests that it may be the cortical area that is responsible for the LVF superiority that was found specifically for edge-based segmentation in the previous experiment.

In experiment 3, therefore, patient DF was tested with the same stimuli used in experiment 1. DF's brain damage is bilateral and is centred on LOC, thus severely impairing her perception of shape. She does, however, retain low-level processes that allow simple segmentation, and she can perform manual actions on objects based on their form, despite those processes not generating visual awareness. With her LOC damage, therefore, she should be capable of segmenting figures based on a single cue-type; if segmentation by a conjunction of features necessitates LOC, however, she should be noticeably impaired at this.

5.7. Experiment 3 – cue-invariant segmentation in patient DF

5.7.1. Methods

Participants

DF's lesion is bilateral and includes complete disruption of LOC (of the ventral stream), whilst her early cortical areas as well her dorsal stream are largely intact. Her remaining areas allow her to guide manual actions (James et al., 2003) and, despite a severe inability to describe object form (Milner et al., 1991), to perform segmentation tasks (Carey, Dijkerman, Murphy, Goodale, & Milner, 2006). DF was 57 years of age at the time of testing. Testing was carried out at the University of Durham's Psychology Department over the course of two days. In addition, four non-age-matched controls took part in a version of the task.

Stimuli and procedure

The stimuli used were the same as in experiment 1 with normal observers, except that the duration of the intermittent mask frame was kept constant at 4 frames, rather than adjusting in line with the frequency of the complete cycle. This was due to DF requiring the stimulus cycle to be substantially slower than that which was required in normal observers to reach a suitable level of performance. At these lower frequency levels, the mask frames would have otherwise been presented for an unnecessarily lengthy duration, during which the target is not present. This would have added a substantial amount of noise in the stimuli and thus an additional degree of difficulty for DF, who already found the task generally a lot more difficult than normal observers. The magnitude of the orientation contrast was also changed; it was increased from 25° to 90°, and the spatial frequency levels were also changed to 1.4 and 3.5 cycles/degree. Additionally, a method of constant stimuli was adopted, and more extensive practice on the task was required for DF. Also, due to time-constraints, only the segmentation conditions were tested. Ideally, given ample time, the discrimination condition would have been conducted in order to show that DF was truly segmenting in the segmentation condition. This would have been extremely problematic, however, due to DF’s visual form agnosia which precludes her from discriminating visual orientation.

Patient DF initially practised on the task with a gradually ascending level of difficulty. Two blocks of 30 trials (10 trials from each of the three conditions) were completed at individual frequency levels, first at 1Hz, then at 2, 3.33, 4.17, 5 and finally 6.25. This was designed to familiarise patient DF with the task as much as possible before determining three levels of difficulty that were likely to sample her accuracy at different levels of her psychometric function. Following this, 9 blocks of 45 trials were completed. Each block contained the 3 contrast types (conjunction, orientation and spatial frequency) at 3 levels of temporal frequency (8, 10 and 12 frames), with 5 repetitions of each.

The control subjects completed a version of this task, but at the 8.3Hz frequency level. This was chosen instead of the values used in DF’s experiment because a preliminary test revealed that control subjects would be performing at ceiling level with these values.

Five blocks of trials were conducted, each containing 20 repetitions of each of the three cue-type conditions.

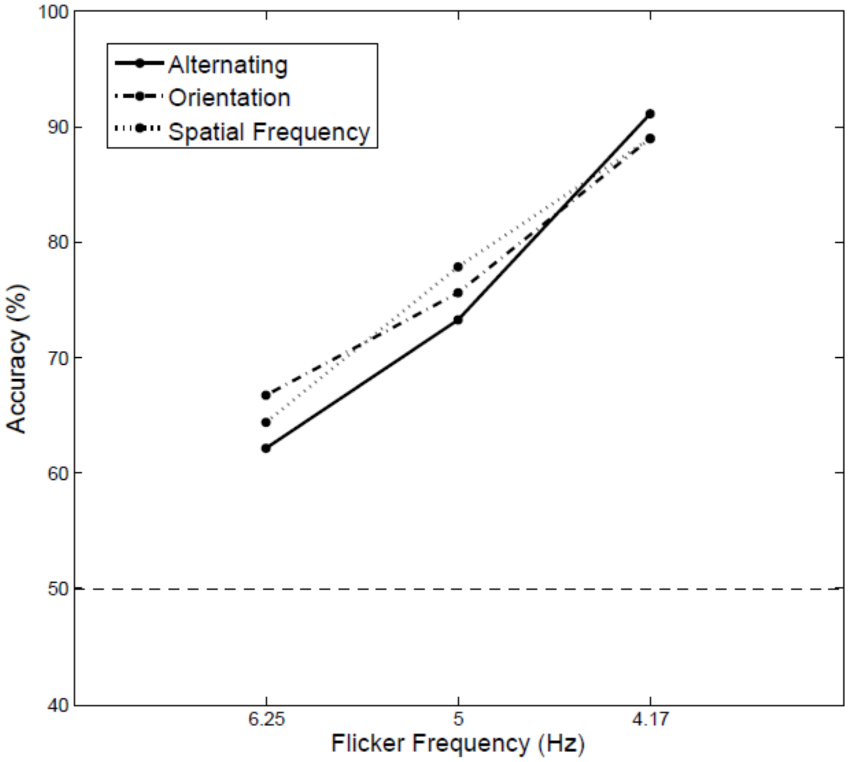

Figure 5.4. Accuracy levels from DF's performance in experiment 3. She did not perform less accurately in the alternating-cue condition than the other two single-cue conditions at any of the frequency levels tested. This suggests that the cue-invariant segmentation stage is to be found earlier than LOC.

5.7.2. Results

Figure 5.4 shows DF's accuracy per cue-type condition across the three frequency levels tested. Note that the frequency levels used sampled her sensitivity across a wide performance range (from ~65% to ~90%). It is apparent that DF's performance was consistent across the three contrast types. Three Chi square analyses were conducted, one for each of the three frequency levels, to assess whether the number of correct responses for the conjunction condition was lower than the other conditions. None were close to significant; for 6.25Hz, $\chi^2_{(2)} = 0.07$, $p>0.95$; for 5Hz $\chi^2_{(2)} = 0.06$, $p>0.95$ and for 4.17Hz $\chi^2_{(2)} = 0.02$, $p>0.95$. This suggests that DF's perception of the figure in the alternating-cue condition was no lesser than that in either of the single-cue conditions. On average, the control subjects scored 88.5%, 89.5% and 89.25% in the orientation, spatial frequency and alternating-cue conditions respectively. A repeated-measures ANOVA was not significant ($F_{(2,6)}$=0.13, p=0.880), indicating no difference in the control subjects' scores between the conditions.

5.8. General discussion of study four

This study explored the visual system's rapid ability to integrate information from different types of 2nd-order cue (orientation and spatial frequency) over time, by using a novel variant of FDF. The technique of FDF was applied in order to dissociate the processes of segmentation from discrimination, as segmentation persists at a higher frequency than discrimination (Rogers-Ramachandran & Ramachandran, 1998; Norman et al, 2011). Experiment one showed clearly that the successful integration of alternating cues for the purpose of figure segmentation occurred much more rapidly than the discrimination of the same regions (i.e. when coincident black border masked the effect of segmentation), and the magnitude of this discrepancy was no smaller for the alternating-cue condition than for either of the single-cue conditions. This ability to integrate information across cue sources is clearly an important aspect of any visual system, as the features that define an object may suddenly change due to environmental changes or due to changes in an observer's perspective.

The feature contrasts used in the present study are low-level computations which are found to occur very early in the hierarchy of the visual system, with evidence indicating areas as early as V1 as potential correlates (Knierim & van Essen, 1992; Sillito, et al, 1995; Lamme, 1995; Hallum *et al*, 2011), and so it is possible to imagine that the single-cue segmentation conditions in the current task could be completed at the level of primary visual cortex. What is unclear, however, is whether segmentation by a rapidly alternating cue type necessitates an additional integration stage beyond primary visual cortex. The second experiment investigated this by comparing segmentation and discrimination of single-cue and alternating-cue types in the UVF and LVF. Upper-lower visual field asymmetries reveal differential processing at the level of V2 and beyond (Horton & Hoyt, 1991; He *et al,* 1996) and so when they are pronounced they are likely to represent higher cortical functions. A general LVF superiority was found in the task; observers performed the task at a higher flicker frequency level in the LVF than in the UVF regardless of whether they were doing the task on the basis of segmentation or discrimination. For segmentation, however, there was a noticeably larger LVF superiority effect in the alternating-cue condition than for either of the single-cue conditions; an interaction which was not present for discrimination of the same stimuli. This result suggests that at least a greater degree of higher cortical involvement was needed for the perception of form in the alternating-cue condition compared to the single-cue condition.

It is important to note, however, that the LVF advantage was not limited to the alternating-cue condition; a significant effect was also found in the single-cue orientation condition. This does corroborate previous results (Graham *et al,* 1999; Hofmann & Hallett, 1993) and suggests that 2nd-order segmentation by a single cue type involves cortical involvement at, or indeed beyond, the level of V2. It could alternatively be, however, a consequence of a visual field asymmetry in attentional resolution (He *et al,* 1996). This is unlikely to be the case, however, as texture segmentation is considered to be preattentive (at least the early components; see Heinrich, Andrés & Bach, 2007), thus not requiring attentional scrutiny. The attention explanation is also not supported by Levine and McAnany (2005), who demonstrated that LVF advantages in visual processing do not vary with task difficulty, indicating that the attention explanation is an unlikely one. Indeed, it is important to note that there is some evidence to suggest that single-cue segmentation for the purposes of form perception may involve some extrastriate areas, namely V2 (Schira, Fahle, Donner, Kraft & Brandt, 2004; Anzai, Peng & van Essen, 2007).

Additionally, the LVF advantage found in the present study was not specific to the segmentation condition; it was found as a main effect in the discrimination condition. It is certainly more likely in this case that attention may be a contributing factor, as the task required observers to judge the stimuli with more scrutiny compared to the segmentation condition. What is important, however, is that the pattern of results found in the segmentation condition was not found in the discrimination condition, implying that the relatively large asymmetry specific to the alternating cue segmentation condition is indicative of some additional involvement of extrastriate cortical area(s). It is also important to discuss whether the visual field asymmetry may in fact simply reflect differential processing within primary visual cortex, and not beyond. It is extremely unlikely that this is the case, as primary visual cortex is divided roughly equally between the two visual fields (Sereno et al, 1995). Consequently, asymmetries at the level of primary visual cortex are rarely reported, and although Liu, Heeger and Carrasco (2006) found stronger activation for LVF stimuli than for UVF stimuli, it was specific only to the vertical meridian. In contrast, the stimuli in experiment two of the present study were presented in visual quadrants (3.26° from the meridian) and so the behavioural asymmetries are unlikely to reflect differential activity solely in primary visual cortex.

The results gathered from patient DF suggest that segmentation by a rapidly alternating cue type is possible without LOC, as she was no less accurate in this condition compared to either of the single-cue segmentation conditions across the three frequency levels included. DF, however, did not show the increased sensitivity in the alternating-cue condition relative to the single-cue conditions, which was present in normal observers in experiment one. The results could be interpreted, therefore, as suggesting that the absence of this multiple-cue facilitation in DF is evidence of LOC's role in cue-invariant segmentation. In the control subjects tested in experiment three, however, no such facilitation was found. This is somewhat at odds with the findings from experiment one, but may be explained by the change in the temporal duration of the intermittent mask frames between the experiments. In experiment one, they were presented for the same duration as the stimulus frames, which varied according to the staircase procedures, whereas in experiment three they were set constantly at 40ms. It may be that the onset-asynchrony between the cue types is an important factor in determining how well the information from different cue sources is integrated. There is no psychophysical work to support or deny this, however, and so it is merely speculation.

Furthermore, as the frequency levels required to assess DF at an approximate threshold level were comparable to those of normal observers when discriminating the regions, it is important to address whether DF was truly "segmenting" the target and not relying on a process of discrimination. As DF's performance was not assessed with the border masks, it is not possible to demonstrate this empirically. It is highly unlikely, however, given that DF retains early segmentation processes (Carey *et al*, 2006), that she was simply discriminating the regions. Additionally, if she were in fact relying on regional discrimination, it would be expected that she would be performing at chance certainly in the orientation cue condition due to her visual agnosia precluding her from discriminating even the most basic orientation signals. As shown by the present results, however, she was clearly able to do this above chance.

The notion that LOC may not be necessary for cue-invariant segmentation is somewhat at odds with previous neuroimaging studies (e.g. Self & Zeki, 2005; Grill-Spector *et al*, 1998) which have shown activity within it to correlate with object perception regardless of the cue from which the object is defined, and even shown increased activity when multiple cues are integrated. The interpretation of the present results, however, is limited to the types of cue used, namely orientation and spatial frequency contrast, and the integration of these cues has not been studied specifically in neuroimaging, making comparisons difficult. The integration of colour and motion, for instance, studied by Self and Zeki (2005) and which is known to be a relatively slow process (Holcomb, 2009), may indeed only manifest at the level of primary visual cortex. Alternatively, however, it may be the case that the observed invariance of LOC only reflects the activity that is fed-forward from preceding areas and does not reflect segmentation *per se*, but rather the construction of an object representation.

In conclusion, it is proposed that the cue-invariant processes that underlie visual segmentation by an alternating cue type are to be found prior to LOC but at least at the cortical level of V2. A reasonable conclusion is that V2 itself is capable of representing cue-invariant contour information, as it is an area known to show signs of cue-invariance when processing contour information (Leventhal *et al,* 1998; Marcar *et al,* 2000).

6. Study Five - Segmenting objects for attention in the absence of awareness

6.1. Aim of the following experiments

This thesis has largely considered the early feature-based processes of the visual system; that is, the elementary computations that are performed relatively early in the visual system that describe 2nd-order contour- and surface-based information. The transition of such information into that of a greater perceptual representation, however, is more complex to understand. The visual brain has at its disposal a number of tools that go beyond these feature-based processes, without which our visual experience would be little more than meaningless lines and colours. What can be termed the *Gestalt* properties of perception describe how the brain constructs such meaningful representations using a set of a *priori* assumptions about the organisation of its external environment. One fundamental task in this regard is the assignment of figure-ground relationships, upon which the contents of our conscious experience is entirely dependent: bistable perceptions, for example, are powerful demonstrations of how scene representations can switch dramatically despite no changes in stimulus information.

How a visual scene is interpreted with respect to its contents is important, as certain perceptual benefits are attributed uniquely to the figures or objects relative to their background (Lamme, 1995), and the edge- and region-based features that are perceived to pertain to a single object are grouped and represented in such a way that attention can be distributed to these features in a single *object*-based manner (objecthood). The results of Naber, Carlson, Verstraten & Einhäuser (2011), for example, quite clearly show that the perceptual benefits that are associated with objecthood (i.e. object-based attention) are not the result of low-level feature-based representations, but of a representation of an object as a “unique representational entity” (Naber *et al*, 2011, p.6). In the stimuli of their experiment, the percept changed spontaneously between that of a bound object and that of an unbound constellation of features, despite no physical changes in the stimulus. Only when the bound object was perceived was there an object-based facilitation effect at and within the object’s borders. The presence of an “object-based attention effect”, therefore, may be taken as an implication of the encoding of an object representation (i.e. a distinct and single perceptual entity that is not restricted to a spatial reference frame). It is unknown, however, at what stage low-level feature information is integrated and encoded into an object representation and, in turn, at what stage this object representation becomes a conscious one. It may be the case that those stages are one and the same; when a perceptual "object" is formed, that object enters consciousness. Alternatively, as it is known that figure-ground relationships are determined early in the visual system (perhaps in primary visual cortex; Lamme, 1995), and may be sufficiently encoded through preattentive processes (Kimchi & Peterson, 2008), an object representation may remain below the level of consciousness and nonetheless be available to influence other processes of perception, attention and action.

The remaining part of this thesis aims to test this theory, by determining whether objects, defined by the rapid 2nd-order segmentation processes documented throughout this thesis, can be selected as units of attention without ever entering awareness. The phase-reversing stimuli that have been used prevalently throughout this thesis offer a novel way of testing this; if an object is defined by an orientation contrast whose sign continually reverses such that the contour is no longer consciously perceived, does the rapid segmentation of the object against its background still allow that object to be selectively enhanced by attention? This section of the thesis will first briefly review what is known of the mechanisms of attention, with focus on the relationship between attention and awareness. The methodological and theoretical limitations of studying attention without awareness will also be discussed.

6.2. Review of attention and attention without awareness

6.2.1. Functional aspects of attention

The functionality of visual attention reflects the need to filter an overwhelming amount of sensory input. As neuronal activity is metabolically and temporally costly, and given the fundamental limits of energy consumption imposed upon the brain (Clarke & Sokoloff, 1994), there is a requirement that cognitive resources are distributed flexibly. This provides the evolutionary bases for a mechanism such as attention to arise in an organism. Some authors claim that it is the very process of attending that turns "looking" into "seeing" (Carrasco, 2011); that is, without a mechanism by which to withdraw from and to engage in the processing of certain information in accordance with task demands or current goals we are effectively blind. Cases of visual neglect, often described as a disorder of attention, certainly bolster this view, as patients suffer from a severe inability to perceive contralesional stimuli.

Inherent within any theory of visual attention is the notion that stimuli in the environment "compete" for neural resources. The photoreceptive cells on the retina encode information about the external world, and once this information is subsequently projected to the subcortical and cortical areas of the visual system, these neural representations of the outside world interactively "compete". The biased-competition model, for example, states that when an organism selectively attends to a particular location in the visual field, the neural representations of information present at that area are enhanced, and those at neighbouring competing regions are suppressed (Desimone & Duncan, 1995); a theory supported by neuroscientific evidence of changes in brain activity (e.g. see Beck & Kastner, 2009). Competition between stimuli has even been observed within the same cortical receptive field (Moran & Desimone, 1985), with the outcome of activity representing a biased average of responses to the individual stimuli (Reynolds, Chelazzi & Desimone, 1999). Although models have been postulated in which the selection criterion is likened to a "spotlight" (Posner, 1980) or a "zoom lens" (Eriksen & Yeh, 1985), such models are not able to fully describe how attention can function by encompassing disparate spatial locations (see Jans, Peters & De Weerd, 2010), or by selecting a unit of perception that is not restricted to a spatial reference frame (this will be discussed later in this section).

It is essential, when discussing the process of attention, to distinguish it from that of *orienting*. Specifically, when attention operates independently of accompanying eye movements it is said to be *covert*. Some see covert attention as being purely concomitant with overt eye movements, and serving only to enhance visual processing at an eventual saccadic destination, thus only being supplementary to overt attention (Findlay, 2005). Others, however, have shown that the two processes have dissimilar characteristics (see Kowler, 2011; Nakayama & Martini, 2011), and one function of covert attention is likely to be the monitoring of incoming information in order to determine the position of upcoming saccades (Carrasco, 2011). Nonetheless, when studying the behavioural and neural effects brought about by attention, it is essential to rule out the possibility that a response enhancement is not solely the consequence of superior processing brought about through foveating.

Covert attention (referred to from here on generally as *attention*) comprises two systems: the endogenous and exogenous systems. William James (1890) described the essential distinction between these systems: the first is voluntary whereas the second is reflexive. In the laboratory, these separate mechanisms are most often studied using various modifications of Posner's (1980) pre-cueing paradigm, the essence of which involves the presentation of a visual target that is immediately preceded by an attentional cue. For endogenous attention, the cue typically takes the form of something symbolic (e.g. a centrally presented arrowhead), and the likelihood with which this stimulus correctly indicates the upcoming target's location is varied. At high likelihoods, an observer is strongly encouraged to utilise the cue, whereas a neutral cue can be selectively ignored (Giordano, McElree & Carrasco, 2009). For exogenous attention, typically a peripheral cue (i.e. a transient luminance flash) is shown before the target, which observers are not able to easily ignore when it is of no informative value (Giordano *et al*, 2009). Observers are quicker, and sometimes more accurate, in identifying/detecting a target when their attention is directed either endogenously or exogenously by a cue, compared to when the cue is not used at all. Conversely, observers are slower when their attention is *mis*directed by a cue. Using such methodologies, voluntary attention is shown to be sustainable over a relatively long time period, whereas the enhancement component of reflexive attention peaks somewhere between 100 and 120ms before the effect disappears and leads to a repulsion effect *away* from the attended region, known as inhibition of return (Posner & Cohen, 1984; Liu, Stevens & Carrasco, 2007; Remington, Johnston & Yantis, 1992). Such behavioural differences, as well as others (e.g. Briand, 1998; Lu & Dosher, 1998, 2000), point to a distinction between purely endogenous and purely exogenous attention, the crux of which is that endogenous attention is under top-down control and is determined by learned associations, whereas exogenous attention is not; it is automatic.

This dichotomy, however, has been challenged on some grounds. Some types of attentional cue, for instance, elicit behavioural effects that are simultaneously suggestive of both endogenous and exogenous processing. The direction of eye gaze in another individual is a powerful cue, for example, even when the cue is non-informative (Friesen & Kingstone, 1998; Driver *et al*, 1999), and induces a rapid facilitation of processing at the cued location (typical of an exogenous system), but also is not followed by inhibition of return (typical of endogenous attention; McKee, Christie & Klein, 2007). Indeed, some have argued for an entirely separate form of attention known as "automated symbolic orienting" to explain such instances (Ristic & Kingston, 2012). Furthermore, when a peripheral cue (i.e. an "exogenous" cue) correctly predicts the location of an upcoming target in a location in space that is non-congruent to that of the cue itself, then facilitation can be seen in this region (Lambert, Naikar, McLachlan & Aitken, 1999). As this involves the learning of a contingency (i.e. a set of rules) to successfully interpret and use the cue, a process which necessitates time to develop in an observer, it is therefore said not to be purely exogenous. Additionally, as such non-congruent spatial contingencies are learned with sub-threshold cues, it casts doubt on another distinction that endogenous attention is conscious and voluntary, and exogenous attention is not. The distinction between what have classically been described as endogenous and exogenous attention, therefore, is certainly not clear-cut; the term exogenous, however, should, if purely for descriptive purposes, be reserved for those instances in which the allocation of attention is automatic (i.e. when the cue's effect is not based on learned contingencies).

Regardless of which form of attention is being discussed, there is a consensus that, generally, attention is a selection process that involves two key components: the *withdrawal* from some things with the *enhancement* of other things. How exactly this is achieved is not exactly known, however; it may be the enhancement of the signal that is the determining factor, or it could be the suppression of the representation of external noise and distractors that are important. It is also possible that attention affects a more cognitive component of perception (e.g. changes in decision-criteria that may involve weighting information from different sources differently). Some authors claim that these may not be mutually exclusive (Carrasco, 2011), and are likely to each contribute to the behavioural effects brought about by attention.

6.2.2. Attention without awareness

The focus of the remainder of this section, and of this thesis, is whether attention can be said to be sufficient for awareness. In other words, is it that the process of attending to something invariably leads to awareness of that very thing? The phenomenon of awareness is not easily understood, but it is clear that it is something which is selective; we can only ever be aware of a fraction of what we potentially could at a given time. The content of awareness, therefore, is determined by some selection process, which has strongly been suggested by some to be attention, implying that what we attend to does, in fact, determine what we see. This echoes the previously mentioned statement at the start of this chapter that attention turns “looking” into “seeing”.

Relying purely on introspection it is hard to disagree with this notion. It is, of course, necessary to test this notion formally and empirically. In theory, one only needs to demonstrate evidence of attention in the absence of awareness to refute this assumption. Unfortunately, this is not as simple as it may have initially been envisioned. As previously discussed, attention operates in many forms and it is often difficult to know exactly how the cue is driving the attentional effect. There are many demonstrations of attention operating without awareness of the attentional cue (e.g. Lambert *et al*, 1999; Kentridge, Heywood & Weskrantz, 1999) and without awareness of the cue’s status as a cue (e.g. Lambert & Sumich, 1996), but it remains most important to show that attention can facilitate the processing of a *target* that remains unseen (Kentridge, 2011). Specifically, one important question that is often asked by many researchers is “Can we attend to something that we are not aware of?” From a philosophical and psychological point of view this is of particular interest because it had long been assumed that prioritization of information by attention was both necessary and sufficient for consciousness (Mole, 2008). Remarkable demonstrations of inattentional blindness, in which otherwise conspicuous visual events are rendered invisible with diverted attention, bolstered this assumption (Mack & Rock, 1998). It has now been shown quite conclusively, however, that attention, at least in the manner envisioned by Posner (1980), is not *sufficient* for generating visual awareness. Kentridge, Heywood & Weiskrantz (1999) demonstrated this in a blindsight patient who declares no awareness of visual experience in his right hemifield as a result of unilateral striate cortex damage and yet retains remarkably accurate performance in some forced-choice discriminations made within that part of his visual field (Weiskrantz, 1986). Selective attentional modulation was observed in this patient in his responses to *cued* stimuli compared with *uncued* stimuli in his blind field, in very much the same way as a normal “aware” observer in a Posner (1980) task. Similar effects have since been found in normal observers, when a masked unseen prime has a greater effect on the subsequent discrimination of a target when attention is directed towards it (Kentridge, Nijboer &

Heywood, 2008; see also Sumner, Tsai, Yu & Nachev, 2006). Arnold, Kinsella and Kentridge (unpublished) have since replicated the original Kentridge *et al* (2008) findings using peripheral exogenous cues in place of central symbolic ones, lending strong support for the claim that attention, regardless of the nature of its deployment, is not sufficient for awareness.

Importantly, these findings cannot be explained by an overt orienting response, as opposed to covert attention, as Kentridge and colleagues monitored eye movements made by subjects throughout the experimental procedures. Furthermore, a general non-specific enhancement (or alerting) response elicited by the cue can not explain the findings, as the blindsight patient studied by Kentridge and colleagues was able to direct his attention to a non-congruent spatial location to that of the cue through the learning of a reverse contingency. It should also be noted at this point that evidence from other fields of cognitive neuroscience support the view of a distinction between the processes of attention and awareness. Magnetoencephalographic recordings, for instance, have pointed to independent neural mechanisms regulating attention and awareness in normal observers (Wyart & Tallon-Baudry, 2008). Stimuli that were consciously seen induced greater mid-frequency gamma-band activity regardless of the subject's attention, whereas the effect of attention on the same stimuli modulated high-frequency activity regardless of the subject's awareness.

In Posner's (1980) initial model of attention, which has exclusively been discussed up until this point, selection is based on a simple spatial primitive in which attention is focussed on a single point in space and spreads uniformly around it. It has been argued that in the cases of dissociation of attention from awareness, it is only *spatial* attention that has been manipulated, whereas awareness has been assessed typically on the basis of the visibility of *objects* (Mole, 2008). In other words, the unit of selection and the object of awareness may not have been truly equivalent in studies claiming to demonstrate dissociations between attention and awareness. Kentridge (2011) argues, however, that, although spatial attention is being allocated, it is in fact the object-related properties (e.g. colour, form) of a stimulus that are being enhanced. This is further supported by evidence that has shown a feature-specific form of attention to be insufficient for awareness (Kanai, Tsuchiya & Verstraten, 2006), and that pop-out based on a feature singleton occurs in the absence of awareness (Hsieh, Colas & Kanwisher, 2011). Attention is, after all, not limited to a purely spatial reference frame; in experiments of Kanai *et al* (2006) and Hsieh *et al* (2011), it is a feature-based model of attention that has been studied, in which attention serves to enhance the processing of spatially disparate stimuli that share a particular feature, at the expense of those that don't share that feature (Carrasco, 2011). Studying such a spatially non-specific form of attention is likely to lead to a more complete understanding of the relationship between attention and awareness more generally, as it allows us to form conclusions on attention irrespective of what is the *unit* of attentional selection.

6.2.3. Object-based attention without awareness?

Of specific interest in the debate of attention without awareness is the phenomenon of object-based attention, in which objects of arbitrary shape form the "units" of attentional selection (Duncan, 1984; Egly, Driver & Rafal, 1994). Egly and colleagues (1994) demonstrated the importance of objects in the deployment of attention using a modification of Posner's (1980) task, in which visual processing of a target was shown to be more rapid when it appeared within the same object as the preceding cue compared with when it was seen to be within a separate object, despite both cue-target pairings being equidistant. This can be explained by a model in which attentional selection operates on the elementary figures that are preattentively segmented by the visual system, although some authors have claimed that attention in this manner is no more than selective spreading of what is fundamentally spatial attention within an object (see Martinez *et al,* 2006). Object-based attention, nonetheless, provides a very good basis on which to test the theory that attention is not sufficient for awareness. This is because the "object" of selection in object-based attention is clear and known, whereas there is some contention as to what is truly selected in spatial attention. Additionally, showing evidence of object-based attention without awareness would further bolster the general claim that attention is not a sufficient precondition for awareness.

Could object-based attention without awareness be possible? Milner and Goodale's (1995) influential framework of the division of labour between the ventral and dorsal streams of visual processing predicts that mediation of object-based attention is to be found in the ventral stream. It is now known more specifically, through neuropsychology, that it is likely to be the lateral occipital complex (LOC), a high level cortical area found on the ventral stream, which is necessary for the effects of object-based attention to emerge (de-Wit, Kentridge & Milner, 2009b). There is much evidence to suggest that activity within the ventral stream is strongly correlated with conscious perception. Single unit recordings in monkeys, for example, suggested that cells within the inferotemporal cortex were only active during conscious perception of a stimulus (Sheinberg & Logothetis, 1997), and there is a considerable amount of neuroimaging work in humans that links the contents of conscious perception to the processes of the ventral stream (e.g. Fang & He, 2005; Pasley, Mayes & Schultz, 2004). Thus, in order to determine whether object-based attention can be found without object awareness we must first ask whether activity in such high-level cortical areas can ever occur without being accompanied by awareness, which would be necessary if we are to assume that object-based attention requires the construction of an object representation in LOC. Some evidence does suggest that activity within the ventral stream is not sufficient for awareness. For instance, a category-specific ERP component indicative of ventral occipitotemporal processing has been shown to be present for face stimuli that never reach awareness (due to binocular suppression; Sterzer, Jalkanen & Rees, 2009). Results from fMRI corroborate this, as areas of the ventral stream have been shown to encode high level object-related aspects of stimulus information (e.g. the presence of a face vs. that of a house; Sterzer, Haynes & Rees, 2008) when the stimulus information is concealed from awareness through binocular suppression. The category-specific activity, however, was only observable through high-resolution fMRI with sensitive multivariate analysis techniques, suggesting that the unconscious activity in high-level areas is weak but nonetheless present. This neural trace may be weak because only a small population of neurons may be capable of encoding category-specific information about stimuli that never reach awareness. Indeed, in the single unit study of Sheingberg and Logothetis (1997), a small proportion of cells in the monkey inferotemporal cortex were stimulus-selective independently of conscious perception despite the overall pattern of responses suggesting otherwise.

As shown, the activity of the ventral stream of processing is not exclusively associated with awareness, as under some conditions it can be shown to be insufficient for generating awareness, albeit with less intensity. The same can be said specifically of LOC; although an object representation specifically within LOC is strongly correlated with the perception of that object (Carlson, Rauschenberger & Verstraten, 2007), it may not necessarily be so (Schira, Fahle, Donner, Kraft & Brandt, 2004). Schira and colleagues, for instance, found that when a figure defined by an orientation contrast was presented outside of the fovea, and when this figure went unnoticed due to attention being directed to a demanding task, neural representation of the figure was nonetheless observed in LOC. Lamme (2003) notes that such activity may be possible in the absence of awareness due to the lack of cortical feedback from the high-level to the low-level areas, which he claims to be necessary for awareness. The initial feedforward sweep activity, he states, is unconscious and can reach parietal and temporal areas within 60ms following stimulus onset. The outstanding question is, of course, whether the unconscious representations of objects found in LOC can be selectively modulated by attention. What is unclear from these experiments alone, however, is whether the object is truly being encoded as an *object* by the brain; without demonstrating behaviourally the advantages conferred by objecthood, this is not possible. Pursuing the effects of object-based attention in the absence of awareness, therefore, would clarify the issue.

In an effort to dissociate object-based attention and object awareness in an Egly *et al* (1994) cueing paradigm, Ariga, Yokosawa and Ogawa (2007) found that awareness of objects was in fact *necessary* to produce object-based attention effects. They used objects defined by illusory contours and found no evidence of a same-object RT advantage when subjects were unaware of the objects. Ariga *et al*, however, manipulated object awareness by changing the preview time of the objects before the onset of the target; in the unaware condition the objects were presented simultaneously with the target, thus there was no delay between the onset of the objects and that of the target. The preview time is important, as objects may need to be present for between 90 and 300ms before the onset of the cue/target before object-based selection develops (de-Wit, Cole, Kentridge & Milner, 2011). Their results, therefore, can simply be explained by the fact that object-based attention did not have a chance to develop in the unaware condition.

Using a different method, with ample preview time, Chou and Yeh (2012) recently showed that object-based attention *could* occur in the absence of object awareness. They masked the objects from awareness using the technique of continuous flash suppression (Tsuchiya and Koch, 2005), in which one eye is presented with a highly dynamic and colourful image to suppress input from the other eye (in this case the objects in an Egly *et al* paradigm) from entering awareness. Two conditions were tested: the *aware* condition and the *unaware* condition, in which the contrast of the rectangles was set at either 100% or 6%.

They found a significant same-object advantage in both cases. For the assessment of awareness they required subjects to discriminate trials in which the objects were present from those in which they were absent in a signal detection procedure. Unfortunately, however, within this test block, both types of conditions were interleaved along with trials in which no objects were present, with only four trials per condition in total. Unsurprisingly, Chou and Yeh found large detection rates (on average 96.25%) for high-contrast objects, but not for low contrast objects (3.75%). As well as there being a potential adaptation-induced after-image subsequent to viewing high-contrast objects, which would have decreased the sensitivity to low-contrast objects (this is certainly the case given that the objects were only ever presented horizontally and therefore in the same retinal location), the authors did not carry out any assessments of sensitivity independently of bias (i.e. a d' measure), and did not conduct nearly enough trials in order to be able to do this confidently. It is impossible to rule out the possibility that subjects simply shifted their criterion along their decision axis in accordance with their being "aware" trials interleaved with "unaware" trials and "absent" trials. In other words, subjects might have been able to detect the low-contrast objects if the high-contrast objects were not included in the block. This point leads fittingly into the following discussion on what is the most appropriate method with which to assess awareness of primes and cues in unconscious attention experiments.

6.3. Assessing awareness of "unseen" stimuli

In tasks demonstrating attention without awareness, it is often very easy to determine whether an effect of attention is present in an experimental setting. What exactly that effect means is perhaps more contentious, but it is more difficult to determine with reasonable confidence that an observer is *not* aware of a particular stimulus. This stems from the inherent difficulty in qualifying awareness, but also from the evidentiary axiom "absence of evidence is not evidence of absence". There are also, however, more methodological problems imposed upon the study of attention without awareness, which are detailed now in this section.

Awareness is an entirely private experience, with no directly observable external effect, so how can it be measured reliably? For many reasons it may seem sufficient to simply ask for a report of consciousness in the observer: "e.g. did you see an elephant?" An advantage to such an approach is that it appeals directly to the *content* of consciousness, but it is lacking in that it will necessitate some memory of the object of consciousness, and numerous biases affect the content of an observer's report independently of their experience. An alternative approach is to employ some procedure that requires the observer to make judgments about some visual information that they may or may not claim to be aware of. This is only valid, of course, on the assumption that observers' responses in such a task are directly and positively related to their visual experience. Such an association is often not found, however, as neurological disorders in visual awareness reveal a counter-intuitive ability of patients to accurately discriminate some property of their visual input whilst remaining adamant that they have no visual experience of such information (Azzopardi & Cowey, 1997; this may also be the case for normal observers - Kolb & Braun, 1995). Thus it will always be questionable if a discrimination performance above chance is found as to whether that truly represents awareness of what is being discriminated. Conversely, a discrimination performance that is not statistically different from chance is considered in some cases to be strong evidence of a lack of awareness (it would be extremely unlikely that an observer may have awareness of an object/event and fail to report it above chance, given that the instructions and task are presented clearly).

In order to minimise the number of erroneous experimental conclusions of awareness (or lack thereof) it is therefore necessary to employ an appropriate design and measurement of discriminability. Most commonly, sensitivity is calculated through comparing an observer's hit rate with their false alarm rate on a signal detection task. In a yes/no design, for instance, observers are presented with a single-interval trial that may or may not contain a target. Observers are instructed to indicate whether the target was present or absent. The table below (**table 6.1**) represents the four outcomes for a given trial in a signal detection task, based on an observer's response relative to the reality of the stimulus.

Table 6.1. Response outcomes in a signal detection task.

	Response: Present	**Response: Absent**
Reality: Present	Hit	Miss
Reality: Absent	False Alarm	Correct Rejection

Thus, an observer’s hit rate (H) is represented by the proportion of trials in which the observer responds “present” for when the target is actually present, or H=p(“present”|PRESENT). The false alarm rate (F), conversely, is represented by the proportion of trials in which the observer responds “present” for when the target is actually absent, or F=p(“present”|ABSENT). Two important statistics can be calculated on the basis of these two values:

- Response criterion (or *bias*): a measure of an observer’s tendency towards one particular response (“Present” or “Absent”) irrespective of the stimulus’s true presence. This is typically represented either as a *likelihood ratio* (**β**) or a *criterion* (**c**). The advantage of using the latter as an estimate is that it is not affected by a change in sensitivity. This is calculated as c=-0.5*[z(H)+z(F)], and is expressed in standard deviation units, representing the distance between the neutral criterion and the observer’s (set) criterion. In other terms, when c=0, an observer does not respond preferentially with either “present” or “absent” (i.e. a neutral criterion). An observer can choose to set a liberal criterion, at which the cost of a miss weighs more heavily than that of a false alarm, or a conservative criterion, at which false alarms weigh more heavily than misses. Thus, the position at which an observer sets their criterion influences the number of “present” and “absent” responses they give overall; this is also known as a *bias*. See **figure 6.1**
- Sensitivity (**d’**): a bias-free measure of an observer’s ability to detect the presence of a target. This is calculated as d' = z(H) - z(F), and is expressed in standard deviation units as the distance between the signal distribution and the noise distribution. A value of 0, therefore, would indicate that the two distributions are indistinguishable for the observer, and is very likely to indicate a lack of awareness to the target. d’ assumes that both distributions are normal and have equal variance. In the more general case, however, d_a comes to represent sensitivity when it is not assumed that the distributions share the same variance:

$$d_a = \frac{(\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2 + \sigma_2^2}{2}}}$$

As this measurement is independent of an observer's response bias (e.g. a tendency to respond "present" consistently more often than "absent" independently of signal sensitivity), d' is a very durable estimation of an observer's sensitivity to the signal.

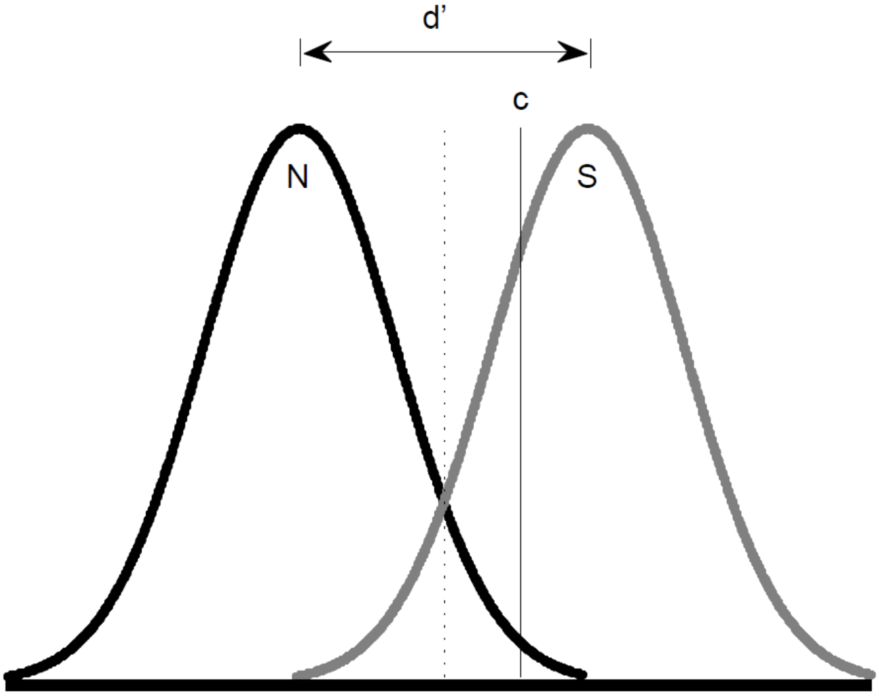

Figure 6.1 Graphical illustration of the decision factors present in signal detection theory, represented on a horizontal perceptual strength axis. The two curves represent probability distributions for neural signals present in an observer for stimuli consisting of *noise* (N; left distribution) and *noise + signal* (S; right distribution). The solid grey line represents a level at which the observer might choose to set their criterion (c), above which they decide *signal,* and below which they decide *noise.* It is assumed in models of signal detection that the two distributions overlap, implying that the two types of stimuli are never truly separable by the observer regardless of their criterion, resulting in misattributing some noise stimuli as signals (false alarms) and some signal stimuli as noise (misses). The dashed vertical line represents a neutral criterion, with no bias in responding with signal or noise. The distance labelled d' describes the observer's sensitivity to the signal, which is what is typically estimated in methods of signal detection theory.

Additionally, it is often desirable to include a confidence-rating scale to accompany observers' responses in a yes/no task. This offers the opportunity for observers to give an indication of awareness on a scale that is not dichotomous. In a yes/no design, for example, an additional confidence rating on a scale of 1-4 provides in total an ordinal scale of 8 responses, allowing a more sensitive measurement of signal sensitivity under the assumption that confidence varies with perceptual strength. Such rating scales allow observers to indicate changes in their bias; an *n*-point scale for instance divides the perceptual strength axis (in **figure 6.1**) into *n* corresponding levels, from which *n-1* criterion points can be inferred. A total of *n-1* points on a receiver operating characteristic (ROC) curve (see **figure 6.2**) can thus be constructed by accumulating the hit rate and false alarm rate as one moves from the most extreme criterion to the lowers criterion. It follows then, that an ROC with a linear plot indicates that across confidence ratings, an observer does not show an ability to systematically minimise false positive rates whilst maximising true positive rates. Conversely, a curve which bows upwards (i.e. large Y values for relatively small X values) betrays the hypothesis that the observer could not distinguish the two conditions and thus could have had awareness of the stimuli.

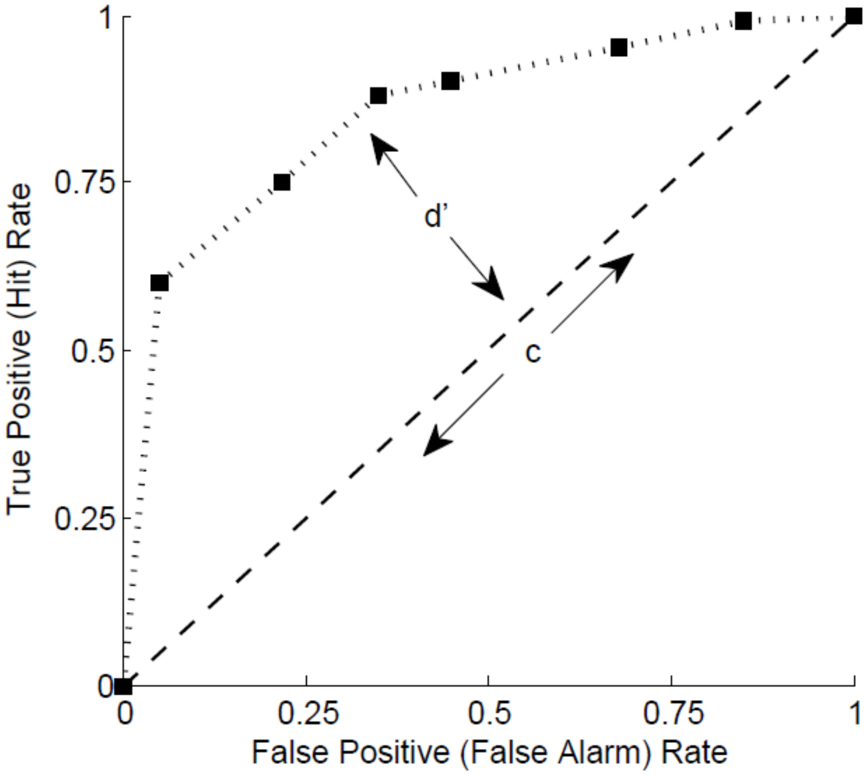

Figure 6.2 An example of a receiver operating characteristics (ROC) curve. The x axis represents an observer's false positive rate (or false alarm rate), and the y axis their true positive rate (or hit rate). The points on the graph show these hit rate and false alarm rates at different levels of an observer's criterion. In psychophysics, these points are usually determined by requiring an observer to give a *rated* confidence response, and so each point comes to represent sensitivity (d') at a different criterion (c).

Thus, an ordinal measurement of sensitivity, using a confidence-rating design, derived from the analysis of hit rates and false alarm rates seems a powerful tool to truly determine whether an observer has any awareness of a stimulus. It should be noted, however, that finding a positive d' with such methods is not entirely sufficient to claim awareness; Azzopardi & Cowey (1997) showed that a blindsight patient (GY) declared no awareness and yet showed above-chance performance in a confidence-rating yes/no task and 2afc task. Thus, problems exist even with carefully constructed signal detection methods, as "good performance" may imply nothing more than the processing of information below a threshold of awareness. Contrarily, a d' measurement that is not significantly different from 0 in a confidence-rating signal detection method is certainly powerful evidence of a lack of awareness. In other words, a d' of 0 may be sufficient but not necessary for implying lack of awareness. A careful approach, therefore, is required that necessitates both a subjective report and an extensive objective measurement of sensitivity.

The validity of the d' measurement, however, is often restricted to the type of task used. This is most applicable to scenarios in which d' is used to claim an absence of awareness of a stimulus (e.g. a prime or cue) which has been shown to indirectly affect perception or performance. An observer may, for instance, show a speeded discrimination of a target that is validly predicted by a masked (unseen) cue, compared to when it is not predicted. A logical step would be then to determine in a subsequent detection task that observers can *not* discriminate the property of the cue that was central to the effect in the previous task (e.g. its location). Some researchers rightfully argue for complete parity wherever possible between tasks that independently measure direct and indirect access to a cue (Reingold & Merikle, 1988). The following criteria have been suggested:

1. Task context: stimulus sequences are presented in the same way
2. Stimulus states: performance of the same task twice, once on the cues and once on the targets
3. Perceptual sensitivity: avoidance of response bias
4. Same response metric: use of same scale to measure effects

Typically, the above criteria are met in studies that set out to determine the unconscious effects of a cue/prime, although often with the exception of the "same response metric". This is because the indirect measurement of the effect of a cue is typically done on a ratio scale (e.g. RT), whereas direct measurement is usually only required on a nominal or ordinal scale (e.g. a dichotomous response, or a confidence-rating scale). In addition to the above criteria, however, Vermeiren and Cleeremans (2012) argue that the two tasks should only be equated insofar as attention is not distributed differently between them. Specifically, if observers are instructed to ignore a cue (or indeed not informed at all of its presence) in a cueing task, and then instructed to discriminate the cue (and only the cue) in a subsequent signal detection task to assess awareness, then the effect of voluntarily directing attention towards the cue results in an overestimation of perceptual sensitivity relative to the true level during the indirect measurement of the cue's effect (which is what was initially intended to be measured!). This is similar to the effects typically observed in studies of inattentional blindness, in which awareness of an otherwise very salient stimulus can be prevented simply by making it irrelevant to the task. Pitts, Martinéz & Hillyard (2012) recently showed in an inattentional blindness paradigm that event-related potentials associated with low-level perceptual processes of contour integration do not automatically determine awareness of the stimuli. This is unsurprising; however, in addition to a later component associated with awareness, there were also further widespread components that dissociated from those of contour integration and awareness and instead seemed to reflect whether the stimuli were *task-relevant.* What could not be concluded in that study, however, was whether the subsequent components associated with task-relevance could impact on processes associated solely with awareness. The isolation of the task-relevant components and awareness components in Pitts and colleagues (2012) study should lead to further clarification on this issue and lead to developing more appropriate measurements of awareness independently of task-relevance.

Thus, there is no consensus on what are the most appropriate methods with which to determine absence of awareness. This is most evident when experiments are designed to dissociate a non-conscious effect of a stimulus (a cue or prime, for example) from its conscious perception. The greatest difficulty arises when attempting to equate the two tasks that independently measure non-conscious and conscious aspects of a stimulus or event, and can often result in overestimating sensitivity. This is important to avoid mislabelling many potentially significant experimental findings as “misses”, when in fact they should be “hits”.

6.4. The present study

The motivation of the experiments in this study was to determine whether objects can act as units of attentional selection even when they are not consciously seen. This would be striking, not only because object-based attention involves a level of sophistication beyond simple spatial selection, but also because there is clear parity between the objects of attention and awareness in the present experiment that may be lacking in tasks solely employing simple spatial selection. In the experiments described here, objects were defined by an orientation contrast to their background, but, crucially, the orientations of the texture elements were continually reversed. In this thesis, it has been shown that orientation-reversing stimuli of this type have previously revealed that the perception of a border between two adjacent regions of texture persists despite the two regions being continually masked (Norman, Heywood & Kentridge, 2011). In the following experiments the orientations were reversed at a frequency above which the conscious perception of the contour also vanishes. A similar phase-reversing masking procedure using luminance-defined words has revealed an unconscious priming effect from masked words that were presented on the display for a total of ~267ms (Bermeitinger, Kuhlmann & Wentura, 2012). This implies that such a technique is capable of masking a stimulus from awareness, and nonetheless eliciting high-level information processing, despite its extended presentation. Additionally, as information that is masked through rapid temporal flicker is nonetheless available to induce attentional orienting (Lu, Cai, Shen, Zhou & Han, 2012), it was hypothesised that objects defined by rapid 2nd-order segmentation processes may nonetheless be processed at a level that allows object-based attention, whereas any awareness of those objects would be prevented by the continual masking of the stimuli. This technique therefore provides a useful tool for masking objects for the length of time that is necessary in a typical Egly *et al* (1994) object-based attention paradigm. If an object-based attention effect is found, it would imply that the objects' contours were being segmented by the visual system's 2nd-order segmentation processes and that information was being integrated into a high-level object representation.

6.5. Experiment 1 – object-based attention without awareness with fixed object positions/orientations

In experiment 1, participants completed a standard Egly *et al* (1994) cueing task, in which they identified the colour of a target that was validly cued (50% of trials), invalidly cued, appearing at a location within the same object as the cue (25%) or invalidly cued appearing in the other object (25%). The objects remained in the same position for the entirety of each block of trials, as in the original Egly *et al* experiment. A subjective assessment of awareness was conducted subsequently to the attention task by asking subjects to answer a series of descriptive questions regarding the content of the experimental display.

6.5.1. Methods

Participants

Sixteen naive observers participated, and all gave their written informed consent. Participants were students recruited through the Durham University Psychology department's participant pool scheme, and were awarded course credits for their participation.

Materials

Stimuli in all experiments were generated using a Cambridge Research Systems ViSaGe Graphics System and were presented on a gamma-corrected ViewSonic 17'' display monitor viewed at a distance of 41cm (participants rested their head on a chin rest). The background had a luminance of 50cdm^{-2}. The screen resolution was set to 1024 x 768 pixels with a refresh rate of 100 Hz. The ViSaGe Graphics System ensured that stimulus display and response timing were time-locked with the monitor's refresh rate.

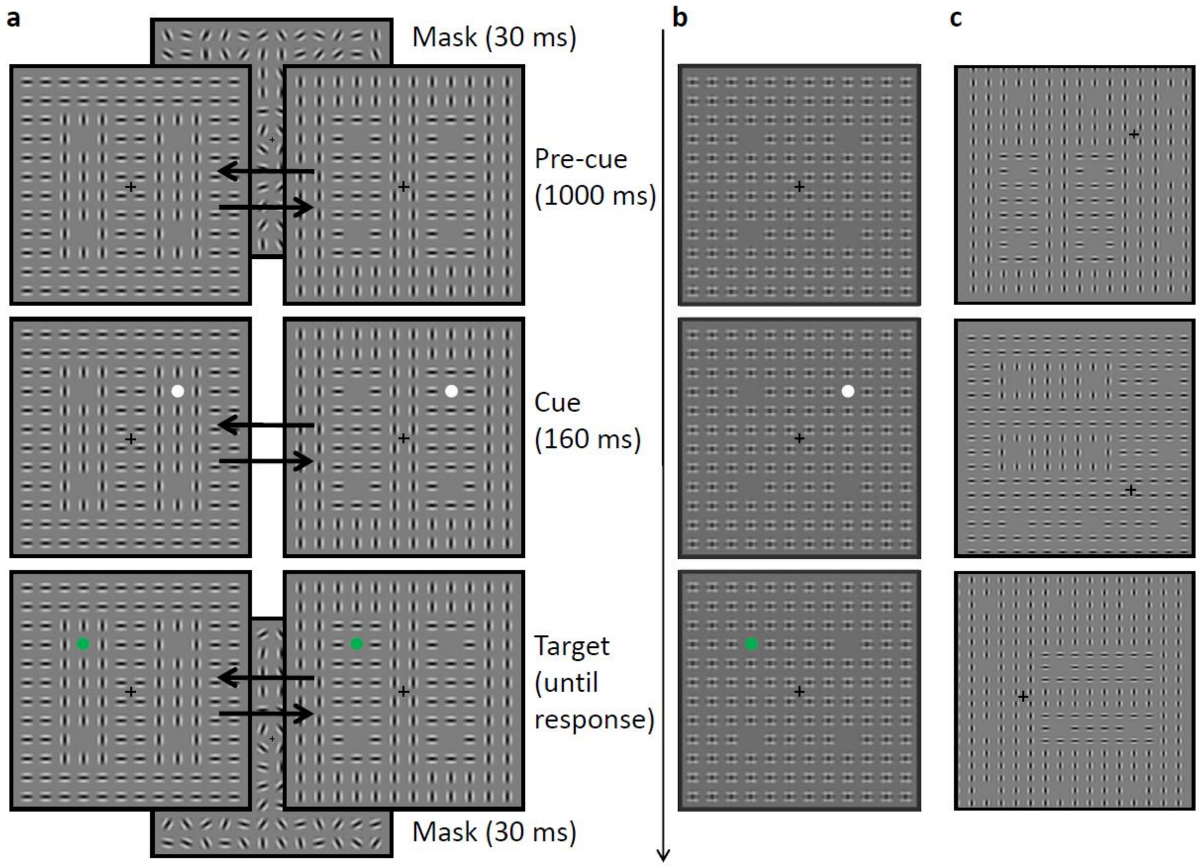

Figure 6.3 Illustrations of the stimuli in experiments 1-3 and of the general temporal sequence for all experiments; a simulated observer's perception of those events; and examples of the different object positions in experiment 4. **a)** Full illustration of the trial sequence of experiments 1-3. Note that there are fewer Gabor patches used to produce the objects in the illustration than in the actual experimental stimuli. Here the target is invalidly cued between objects. Double arrows indicate that the two frames are presented continually in alternation at a frequency of 16.7 Hz. The temporal sequence described here can be applied generally to the other experiments. **b)** A simulated observers perception of the sequence shown in figure **a**. Participants are not aware of the presence of the figures. **c)** Examples of three of the eight object positions used in experiment 4; relative to the fixation cross, they are bottom-left, top-left and centre-right. Note that the illustration represents only a magnified section of the full stimuli. The cue and target on each trial would only be presented within the 4 placeholders associated with the figures.

Procedure

Participants fixated a central cross. Following a warning tone, a lattice (21° in width and in height) centred on the fixation cross consisting of 18 x 18 uniformly positioned Gabor patches appeared. Each Gabor had a diameter of 0.6° and a spatial frequency of 2.7 cycles/degree and was separated from its neighbours by 0.6°. Each Gabor had a maximum contrast of 100% with a Gaussian standard deviation of 0.2°. See **figure 6.3a** for illustrations of the stimuli and procedure. For 30ms, these Gabors were presented, each with a randomly determined orientation, as a mask before the onset of the objects. Immediately following this, the Gabors would continually alternate between vertical and horizontal orientations at 16.7Hz. Two identical rectangular objects (measuring 12 x 3 Gabor patches and presented 2 Gabors either side of fixation) were formed by an orientation contrast of 90° to the background. Thus, the objects were always defined by an orientation contrast of 90° to that of their background even as the orientations of *all* the Gabors in the display alternated. In one session of trials they were positioned horizontally, above and below fixation, and in another they were vertical, to the right and to the left. In four of the positions in the lattice, located 6 Gabor spaces vertically and horizontally in from the four corners, no Gabors were presented, as these locations served as placeholders for cues or targets. Thus, one placeholder was located at either end of both figures.

1000ms after the onset of object presentation a cue (white disc (luminance = 158 cdm^{-2}), 0.4° in diameter) appeared for 160ms in one of the four placeholders (determined randomly with equal probability on each trial). Following the offset of the cue, the target disc (0.4° in diameter) appeared in one of three locations, and would either be red (CIE 1931 x, y coordinates of 0.40, 0.31 with a luminance of 72.14 cdm^{-2}) or green (CIE 1931 x, y coordinates of 0.30, 0.59 with a luminance of 81.06 cdm^{-2}). In valid trials (50% of all trials) the target would appear in the same position as the cue. In invalid-within trials (25%) the target appeared in the adjacent placeholder that was within the same figure as the cue had been. In invalid-between trials (25%) the target appeared in the adjacent placeholder that was within a different figure. The colour of the target was determined randomly with equal probability on each trial. Participants were instructed to discriminate the colour of the target disc by pressing one of two buttons. The target remained on the screen until a response was made; following which the noise mask of random orientations was presented again for a further 30ms, ending the trial. See **figure 6.3** (**a** and **b**) for a depiction of the display sequence and a simulation of the observers' perception during the sequence, respectively.

Participants completed two blocks of 140 trials, each with a different figure orientation, with the order being counterbalanced across participants. Following the completion of both trial sessions, participants were asked a short series of questions designed to assess whether they had any awareness of the presence of figures in the display. These were, in order: Can you identify the purpose of the task? Can you describe what you saw on the screen? Did you see anything other than the cross, flickering lines, white flash *(cue)* and target? Did you notice any outlines of shapes in the background of flickering lines? Was there any difference between the first and second session *(one contained vertical figures, the other horizontal)*? After answering the questions, participants were then given a few additional trials with the reduced alternation rate of 4 Hz and were asked to describe anything in the display that was not visible in the experimental trials.

6.5.2. Results and discussion

Only correct trials were analysed. The RTs were trimmed by first removing those that exceeded 1500ms or were less than 150ms, interpreted as unsuccessful button-presses or anticipatory responses. The remaining data that fell outside 2 standard deviations from the mean per condition per participant were removed as outliers. This procedure removed 9.3% of all trials. A within-participant ANOVA with the single factor Cue Validity was conducted on the mean values of the remaining RTs, where the overall means were 407.8ms (valid), 426.0ms (invalid-within) and 432.7ms (invalid-between) as shown in **figure 6.4**. The main effect was significant ($F_{(2,30)}=82.73$, $p<0.001$) indicating that the cue had a different effect on participants' RTs depending on its position relative to the target and the figures. A planned paired t-test revealed that RTs were significantly shorter on invalid-within than for invalid-between trials ($t_{(15)}= 3.68$, $p=0.002$), indicating that participants were quicker to respond to targets that appeared within the same object as the preceding cue relative to those that appeared in a different object; a classic replication of the object-based attentional effects of Egly and colleagues (1994). No significant effect of accuracy was found between the conditions of cue validity ($F<1$); this indicates no trade-off between RT and accuracy.

In response to debrief questions following the experimental sessions, participants showed no indication that they had any awareness of the objects that had influenced their responses. When asked to speculate as to what the experiment's aim might have been, most participants made some general comment on cue validity or congruency between target position (left/right) and response required (left/right), but none mentioned anything related to the presence of objects in the display. Similarly, when asked to describe what they saw on the screen they made no mention of the objects and none could identify that one session contained horizontal objects and the other vertical. Indeed, when the participants were given the opportunity to view the display at a slower speed and asked to describe anything that had previously been invisible to them, they immediately pointed to the objects, accurately describing them as two rectangles. This indicates that participants' RTs, and thus their attention, were influenced independently of their awareness of those very objects that were responsible for the effect.

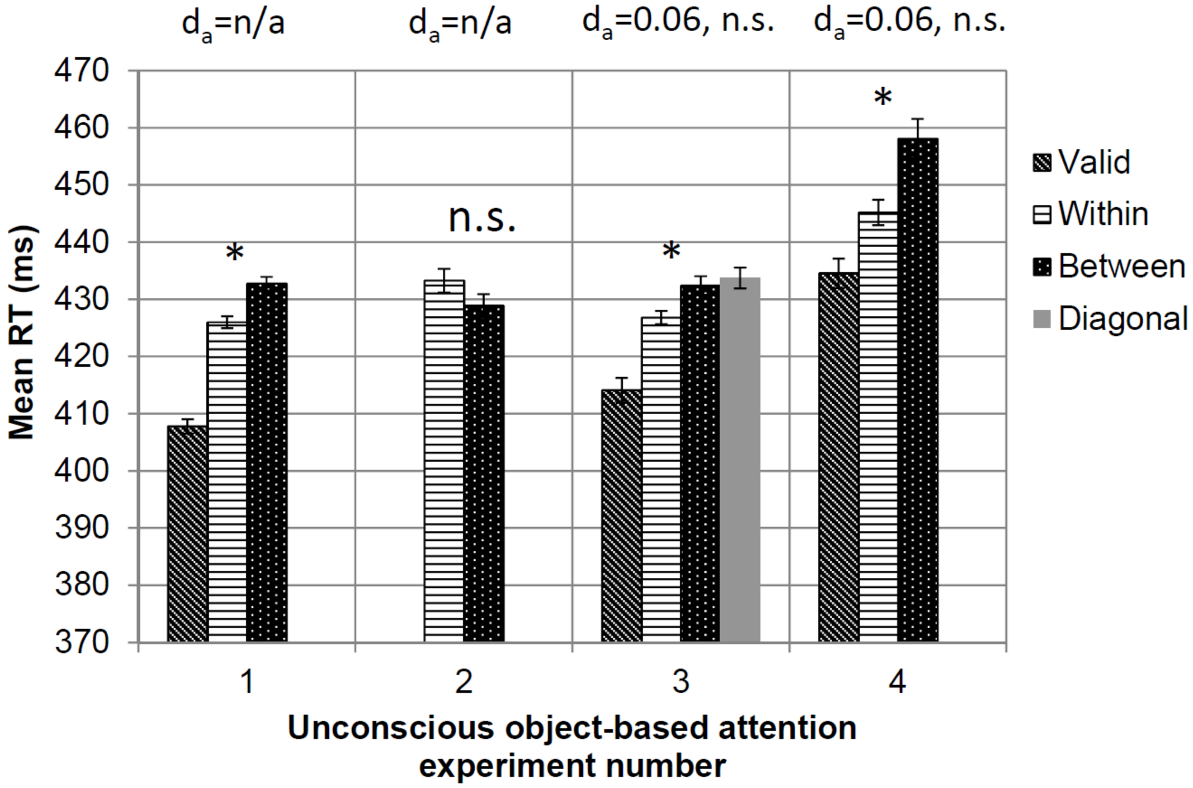

Figure 6.4 RT results from unconscious object-based attention experiments 1-4, along with associated d_a values and statistics (where applicable). Results from experiment 1 (N=16), 2 (N=9), 3 (N=20) and 4 (N=8). Participants are quicker to identify the colour of a target when it appears within the same object as a preceding cue compared with when it appears on a different object, whilst spatial distance is equated. This is true despite participants denying any awareness of those objects. In experiment 2, however, in which only within- and between-object conditions were included, this was not the case. Single asterisks denote a significant paired t-test at the 5% level; double asterisks denote the same at the 1% level. Error bars show 1 standard error of the mean with between-participant variance omitted**.** Experiment d_a values and statistics are displayed above the associated experiment (for 3 and 4); neither of the values were significantly above zero.

Although there is evidence here that object-based attention may operate without awareness, it can be argued that, despite the targets appearing at the two invalidly cued locations with equal probability, there is, nevertheless, a contingency which might drive implicit learning of an object-based selection strategy. Targets appear three times as often within the same object as the cue (50% valid plus 25% within object invalid conditions) as they do in the other object (25% between objects invalid condition). An important aspect of object-based attention, however, is that it should be determined simply by the arrangement of objects in the scene, rather than as a consequence of some implicitly learned association; it should be exogenous. Lee, Kramer Mozer & Vecera (2012) manipulated cue-object contingency and showed that when the target appeared equally likely on the cued and non-cued object, thus nulling the contingency, subjects were still overall quicker to respond in the within condition than in the between condition. This demonstrates the exogenous nature of object-based attention, which is how it is typically assumed to operate. It is unclear, therefore, whether the effect observed in experiment 1 of this study is the result of the automatic processing of the objects, or because of the object-based contingency between the cue and target. Although it has not been shown conclusively, however, that such an effect could indeed arise purely from the presence of a contingency, it *has* been shown (Lee *et* al, 2012; Yeari & Goldsmith, 2010) that object-based attention may only be a default setting by the visual system which can be overridden by a highly competing spatial strategy; in other words, participants could restrict their attention to only *one* part of a perceptually completed object when it was highly strategically viable to do so. The aim of the next experiment was thus to determine whether, in normal, conscious object-based attention, an object-based contingency is sufficient to produce an object-based attention effect.

6.6. The importance of contingencies in (conscious) object-based attention studies – an experiment

The aim of this experiment was to demonstrate how an object-based attention effect can arise purely from a contingency between the cued-object and the target, and not because of the parallel processing of the objects in a scene. To show this, a transient cue is used which predicts the target to appear more often on the non-cued object in a standard

Egly *et al* task. This contingency is created by retaining the within- and between-object trials (each for 20% of trials), but removing the valid trials and replacing them with “diagonal” trials (60% of trials), in which the target appears in the diagonally opposite placeholder to that of the cue, and thus on the opposite object. If an object-based attention effect can be brought about through the learning that the target appears more often on a particular object relative to the cue (in this case, the “non-cued object”; in the typical case, the “cued” object), then RTs for target identification should be shorter for “between” object trials relative to “within” object trials.

6.6.1. Methods

Participants

Twenty naïve subjects took part in this experiment.

Materials

See previous experiment.

Procedure

The general procedure was the same as in the previous experiment, except that the objects were visible to the observers (i.e. the orientations of the Gabor patches did not alternate). In addition, the SOA between the onset of the objects and of the cue was 500ms, and the objects could assume either vertical or horizontal orientations (distributed randomly throughout each session).

Subjects completed five sessions; each session contained 150 trials. Within each session, three cue-target conditions three cueing conditions were included. In “diagonal” trials (60% of all trials) the target would appear in the diagonally opposite placeholder as the cue. In “within” trials (20%) the target appeared in the adjacent placeholder that was within the same figure as the cue had been. In “between” trials (20%) the target appeared in the adjacent placeholder that was within a different figure.

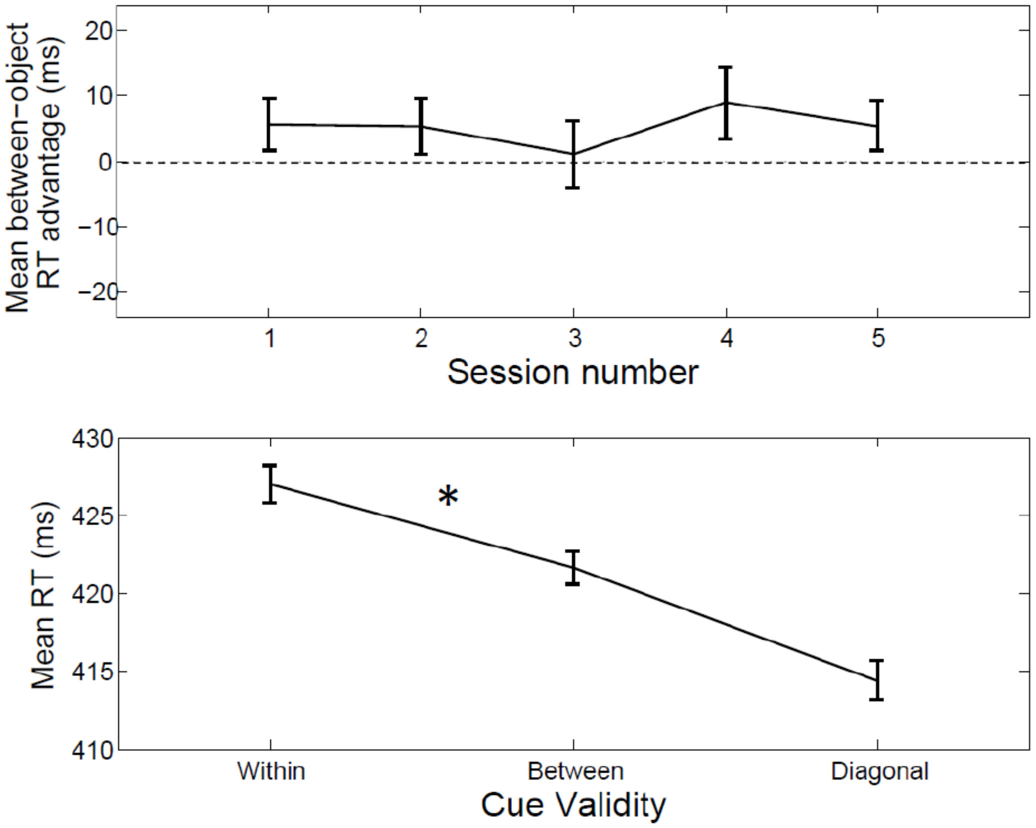

Figure 6.5 RT results from the experiment demonstrating the importance of the cue-object contingency in (conscious) object-based attention experiments. The lower figure shows that, on average, subjects identified the colour of the target more quickly when it fell on a different object to that of the preceding cue. This is because the cue predicted that the target would appear on the non-cued object 80% of the time. The upper figure shows that this between-object RT advantage is apparent from the very first session, indicating that it is not a contingency that requires excessive exposure to be learned and to influence behaviour. Single asterisks denote a significant paired t-test at the 5% level; double asterisks denote the same at the 1% level. Error bars show 1 standard error of the mean with between-participant variance omitted**.**

6.6.2. Results and discussion

RTs were trimmed in the same way as previously described, removing 10.4% of all trials. The remaining RTs were averaged across the sessions, and these averages were entered into a repeated measures ANOVA, with the single factor of cue validity (within, between, diagonal). This was significant ($F_{(2,38)}$=18.49, p<0.001), indicating that subjects’ RTs were affected by the validity of the cue. The average RTs for each condition were as follows: 426.9ms (within), 421.7ms (between) and 414.4ms (diagonal). Crucially, a planned paired t-test revealed that subjects were overall quicker to respond in the between-object condition than in the within-object condition ($t_{(19)}$=2.79, p=0.012). See **figure 6.5** for the results. The finding that the diagonal location of the target yielded the shortest RT of all three conditions despite being the furthest distance from the cue is not surprising, as it is known from previous research that RTs to a target that appears in a spatially non-congruent location relative to the cue are relatively short provided that the location is validly predicted by the cue (Lambert *et al*, 1999).

Accuracy rates were also analysed with a repeated measures ANOVA with the single factor of cue validity. No effect was found ($F_{(2,38)=}$0.89, p=0.418), indicating no trade-off between RT and accuracy across the 3 conditions.

Further statistics were conducted to ensure that the observed between-object advantage in RT did not arise solely from an effect of the cue-validity on the immediately preceding trial, and instead reflected the contingency that could only have been learned through exposure to many trials (see Lee *et* al, 2012). Criteria for outlier removal and RT analysis were similar to previously described, although trials were organised with respect to two factors: current trial validity (within vs. between) and previous trial validity (within vs. between). The factor previous trial validity refers to the validity of the cue on the immediately preceding trial, in which “between” pertains to both the between and diagonal cue-validity conditions, as in both cases the cue is followed by the target on a different object. As the previous trial validity must be included as a factor for this analysis, the first trial from each session was excluded from analysis by necessity. The repeated measures ANOVA revealed a significant main effect of current trial validity ($F_{(1,19)}$=5.06, p=0.037), with between-object RTs being overall shorter (424.9ms) than within-object RTs (431.0ms), and a marginally non-significant main effect of previous trial validity ($F_{(1,19)=}$4.11, p=0.057), with RTs being shorter for trials following between-object trials (424.8ms) than those following within-object trials (431.1ms). Importantly, however, there was no evidence of an interaction ($F_{(1,19)}$=0.19, p=0.670), indicating that the between-object RT advantage was not specific to between-object trials that were immediately preceded by a between-object trial. These results further support the notion that the significant between-object advantage arose through the cue-object contingency that was present throughout the experiment, and not through a more transient contingency that may have been found on a trial-by-trial basis.

This experiment has therefore shown that the typical object-based attention effect that is reported in many studies, and which is claimed to be exogenous, is not necessarily so.

The “between-object” RT advantage shown here can only have arisen by subjects directing their attention to the non-cued object after learning that the target is more likely (80% of the time) to appear on that object. This is yet another experiment that challenges the dichotomy of purely exogenous and endogenously controlled attention (e.g. Lambert *et al*, 1999), as the processing of the transient, peripheral cue in this experiment, and the “automatic” segmentation of the objects, is sensitive to the contingencies of the experimental procedure. Importantly, as seen in **figure 6.5**, the effect of this contingency is apparent in the first session, indicating that it can be learned and used by an observer in little time (fewer than 150 trials). Any experiments aiming to demonstrate the effects of “exogenous” object-based attention, therefore, including the first experiment described in this section, should consider the effect of the contingency between the cued-object and the target. It is important, therefore, to determine that the object-based effect found in the first experiment of the present study is not the result of a learned contingency between the object that is predicted by the cue and the target.

6.7. Experiment 2 – removing object-based contingencies

The purpose of this experiment was to show that the object-based attention effect observed in experiment one was purely exogenous and not the result of an unintended contingency. In order to null the effect of this contingency, the orientation of the objects was randomly determined on each trial and the valid cue-target validity condition was removed, leaving only the within- and between-object conditions. If objects can be selected as units of attention in the absence of awareness, then a within-object RT advantage should nonetheless be found with the use of this neutral cue.

6.7.1. Methods

Participants

Nine new naïve participants took part. They were recruited in the same way as previously described.

Materials

See experiment 1.

Procedure

The methods were very similar to experiment one, except that one session of 300 trials was used. Within the session, an equal number of within-object and between-object trials were presented, but no valid trials. An equal number of trials contained vertical or horizontal objects, and these were randomly distributed throughout the session.

6.7.2. Results and discussion

RTs were trimmed in the same way as previously described, removing 8.6% trials in total. A paired t-test was conducted between RTs for the within- (433.3ms) and between- (428.89ms) object conditions, as shown in **figure 6.4**. There was no significant effect ($t_{(8)}=1.073$, $p=0.315$). Additionally, no significant effect of accuracy was found between the conditions ($t<1$). As in the previous experiment, subjects could not describe the objects when prompted to.

Despite previous findings showing that object-based attention can be elicited through a purely exogenous, transient cue (Yeari & Goldsmith, 2010), no such effect was found in the present experiment. There is, of course, an important methodological difference between the two studies: in the Yeari and Goldsmith (2010) experiment, the cue was neutral with respect to the objects because the target appeared equally likely in any of the four placeholders regardless of where the cue appeared. In the present experiment, however, only the standard within- and between-object conditions were included, thus the target never appeared in the same location as the cue, whereas it did so 25% of the time in Yeari and Goldsmith's (2010) experiment.

6.8. Experiment 3 – a second approach to removing object-based contingencies and introducing an objective assessment of awareness

The purpose of this next experiment was again to determine whether purely exogenous object-based attention can occur without awareness, by using the four cue-target conditions of Yeari and Goldsmith (2010). Targets therefore appeared in each of the four possible locations with equal probability (25%) and so were equally likely to appear in the same object as the cue or in the other object. Additionally, this experiment improves on the previous ones as it includes an objective measure of awareness- a signal detection procedure following the attention task. There are many things to consider when devising a measure of awareness, and it is often difficult to achieve a measure which does not result in an overestimation of sensitivity (Vermieren & Cleeremans, 2012). An objective measure of awareness is arguably more preferable to that of a subjective measure, however, because it can represent an observer's sensitivity to a particular stimulus (relative to noise) independently of any response bias they might have. Additionally, when assessing awareness using subjective measures (e.g. through post-experiment debriefs), it can be largely dependent on the subject's memory of the task and stimuli as well as a response bias. This experiment, therefore, also aims to show evidence of an object-based attention effect along with objective evidence that these objects are not consciously accessible.

6.8.1. Methods

Participants

Twenty new naïve participants took part. They were recruited in the same way as previously described.

Materials

See experiment 1.

Procedure

The general procedure remained the same as in experiment 1 and 2, and throughout each block of trials the orientations of the objects (horizontal/vertical) varied randomly from trial to trial. In addition, four target locations were used, the fourth location being in a placeholder in the non-cued-object diagonally opposite the cue location. Targets appeared at each of the four locations with equal likelihood, ensuring that there was no contingency between the cue and the object within which the target appeared.

They each completed one practice block followed by 4 test blocks of 96 trials (48 "horizontal" and 48 "vertical", randomly distributed). Following the completion of all trial sessions and, as before, participants were asked the series of questions designed to probe awareness of the presence of figures in the display. In addition, however, subjects also completed a signal detection task in which they were required to explicitly detect the presence of the objects. Subjects completed two blocks, of 140 trials each, in which on each trial they were presented with the flickering background which contained the objects on half the trials. These trials were randomly interleaved with those in which the background was completely uniform (i.e. all of the orientations in the display simultaneously alternated between horizontal and vertical). For those trials in which the objects were present, there were an equal number of horizontal and vertical trials. The display duration was 1660ms, equivalent to the duration of the stimuli up until the onset of the target in the attention task, after which the subject indicated with one of two buttons whether the objects were present or absent, and then rated their confidence with that judgment on a scale of 1-4.

6.8.2. Results and discussion

Data from one of the participants were not included in the analysis because the goodness of fit for curve fitting in the d_a estimation was significantly low ($p=0.018$). The RTs from the 4 test blocks were trimmed as before, resulting in exclusion of 7.5% of trials. A repeated measures ANOVA revealed a significant main effect of cue validity on RT ($F_{(3,54)}=18.23$, $p<0.001$), where the overall means were 414.1ms (valid), 426.8ms (invalid-within), 432.4ms (invalid-between) and 433.7ms (invalid-diagonal), as shown in **figure 6.4**. A planned paired t-test determined that RTs were significantly shorter for invalid-within trials compared with invalid-between trials ($t_{(18)}=2.59$, $p=0.019$). No significant effect of accuracy was found between the conditions of cue validity ($F<1$); this indicates no trade-off between RT and accuracy.

As in the previous experiment, subjects could not describe the objects when prompted to. Data from the signal-detection task were used to measure participants' sensitivity to the presence of the objects. Sensitivity was calculated by tabulating the number of responses for each of the eight confidence levels (4 ratings for both "present" and "absent" responses) for both "present" and "absent" trials. The discriminability index d_a was calculated from these data using the software RScorePlus (Harvey, 2002) to fit a Gaussian unequal variance model. d_a allows for unequal variance and is equivalent to d' in the case of equal variance. A higher d_a indicates a greater sensitivity to the target, and a d_a of zero indicates no sensitivity. Overall, the average d_a was low (0.06), and it was not significantly above zero ($t_{(18)} = 1.88$, $p = 0.077$), indicating that participants could not discriminate the signal (objects) from the noise (no objects).

This experiment again demonstrates that object-based attention can occur in the absence of awareness. Importantly, because there was no contingency between the cued-object and the location of the target, this was a purely exogenous from of object-based attention. Furthermore, as the magnitudes of the within-object advantages found in this experiment and experiment 1 are comparable in size, the contingency between the cued-object and target in experiment 1 did not have a considerable impact on the behavioural effects of object-based attention in the current paradigm. Additionally, the objective assessment of awareness in addition to the subjects' subjective reports strengthens the claim that subjects were not aware of the objects. The average sensitivity rating was, however, only marginally not significantly different from 0, casting doubt on whether awareness of the objects was truly and entirely absent, and instead suggesting that perhaps awareness was only attenuated to a very low level. This may have emerged as, in the current experimental setup, subjects could in theory have performed the detection task by scrutinising only a small single part of the display (for instance, at the location of an orientation contrast between the background and objects), rather than awareness of the objects *per se*.

6.9. Experiment 4 – object-based attention without awareness with unpredictable object positions and orientations

As the previous experiment found only a marginally non-significant d_a score, it is important to show that the attention effect is still found when there is more conclusive evidence for absence of awareness. In order to pursue this, the subsequent experiments used a modified version of the stimuli in both the attention and awareness tasks. Specifically, the objects are made smaller and can appear unpredictably in a number of locations throughout the display. An important aspect of object-based attention, in common with exogenous spatial attention, is that it can be deployed rapidly and in parallel across the visual scene (de-Wit, Cole, Kentridge & Milner, 2011). The unpredictable placement of the objects, therefore, should have no significant effect on the observed object-based attention effect. As subjects cannot perform the signal detection task under these conditions by scrutinising only a single small area of the display, their ability to detect the presence of the objects may be shown more conclusively to be no different than chance.

The objects in the present study therefore appeared at unpredictable locations and assumed unpredictable orientations on a given trial. This is important also because it appeals to the automatic nature with which object-based attention is thought to take place.

6.9.1. Methods

Participants

Eight new naïve participants took part. They were recruited in the same way as previously described.

Materials

See experiment 1

Procedure

The general methods remained the same as those in experiment 1; however, the objects were smaller and could occupy one of 8 potential positions on each trial and could be either vertically or horizontally oriented. The background lattice was thus extended to consist of 22 x 22 Gabor patches (diameter of 0.4° and spatial frequency of 3.75 cycles/degree) and contained 16 cue/target placeholders distributed in a 4 x 4 arrangement centred on the fixation cross with each one separated by 4 Gabors from its neighbours (vertically and horizontally). Thus, there were 8 locations where the objects could be defined: top-left, top-centre, top-right, centre-left, centre-right, bottom-left, bottom-centre and bottom-right. The figures measured 8 x 3 Gabor patches separated from one another by 2 Gabor spaces. See **figure 6.3c** for an illustration of the stimuli. The location of the pair of figures would be randomly determined with equal probability on each trial. As in experiment 1, three target locations were used; valid (50% of trials), invalid-within (25%) and invalid-between (25%). The temporal sequence of the experiment was identical to that of the first.

Eight new naive observers participated, and all gave their written informed consent. Participants completed one practice block followed by two blocks of 160 trials, each containing 80 “vertical” and 80 “horizontal” trials distributed randomly across each block. In place of the post-experiment questionnaire, which was not used, a confidence-rating signal-detection procedure was carried out for each participant. Participants were presented with the flickering background from the attention task with or without the objects present in one of the 8 locations in the display. When present, the locations and orientations of the objects would be determined randomly with equal probability. Participants had to indicate on which trials the objects were present and on which they were absent, by pressing one of two keys, and then to rate their confidence with that judgment on a scale of 1-4, by pressing one of four keys. Participants were presented with 300 trials, consisting of 150 “present” and 150 “absent” trials, randomly distributed. The stimuli would remain onscreen for 1160 ms (equivalent to the pre-cue + cue stages of the attention task) before subjects were prompted to respond by an auditory cue. Prior to this task, participants were not debriefed as to the exact nature of the objects; they were instead only instructed that on some trials the flickering pattern would be uniform throughout and that on others the display would contain a non-uniformity occurring at various unpredictable locations in the pattern. Participants were not shown the exact nature of the objects in order to better reflect their expectancies during the attention task, in which they had no knowledge of the structure of the pattern or of the objects that were driving the attentional effect.

6.9.2. Results and discussion

RTs from the attention task were trimmed with the same method as the previous experiments, removing 7.9% of all trials. A within-participant ANOVA with respect to the single factor Cue Validity was conducted on the mean values of the remaining RTs, with overall means of 434.5ms (valid), 445.2ms (invalid-within) and 458.1 ms (invalid-between) as shown in **figure 6.4**. The main effect was significant ($F_{(2,14)}$=11.78, p=0.001) indicating that the cue had a different effect on participants’ RTs depending on its position relative to the target and the figures. A paired t-test revealed that RTs were shorter on invalid-within than for invalid-between trials ($t_{(7)}$= 2.48, p=0.042), indicating that participants were faster to respond to targets that appeared within the same object as the preceding cue relative to those that appeared on a different object. No significant effect of accuracy was found between the conditions of cue validity (F<1); this indicates no trade-off between RT and accuracy.

Data from the signal-detection task were used to measure participants’ sensitivity to the presence of the objects. Sensitivity was calculated by tabulating the number of responses for each of the eight confidence levels (4 ratings for both “present” and “absent” responses) for both "present" and "absent" trials. The discriminability index d_a was calculated from these data using the software RScorePlus (Harvey, 2002) to fit a Gaussian unequal variance model. d_a allows for unequal variance and is equivalent to d' in the case of equal variance. A higher d_a indicates a greater sensitivity to the target, and a d_a of zero indicates no sensitivity. Overall, participants' d_a did not differ significantly from zero (mean $d_a = 0.04$; $t_{(7)} = 0.8$, $p = 0.476$), indicating that participants could not discriminate the signal (objects) from the noise (no objects), and hence could have had no awareness of the objects.

This experiment has shown that object-based attention occurs without awareness when the objects appear in unpredictable locations and assume unpredictable orientations on each trial. This finding is consistent with how conscious object-based attention is thought to operate (in parallel across a visual scene) and so reinforces the notion that the effect demonstrated in these experiments truly reflects the workings of an "ordinary" object-based attention mechanism. Furthermore, the statistic from the signal detection experiment is more conclusive in showing absence of awareness of the objects, further strengthening the claim that awareness of the objects was truly absent

6.10. Experiment 5 – assessing awareness whilst maintaining parity between the attention and detection tasks

In the previous task, evidence of object-based attention was found when subjects could not detect the presence of the objects significantly above chance in a signal detection procedure. Criticisms could be made, however, on the basis that in the signal detection procedure, the exact nature of the objects was not revealed to the subjects. The decision not to reveal the nature of the objects was made in order to equate the subjects' knowledge between the two tasks as much as possible and thus avoiding an overestimation of sensitivity. It could be argued, however, that the non-significant d_a may not represent a lack of awareness, but rather the inaccurate completion of the task by the subjects. For the following experiment, therefore, it was decided to equate the tasks as much as possible *and* to reveal the nature of the objects to the subjects.

Participants completed a standard Egly *et al* (1994) cueing task, in which they identified the colour of a target that was validly cued (50% of trials), invalidly cued within the same object as the cue (25%) or invalidly cued in another object (25%). The objects appeared at unpredictable locations and assumed unpredictable orientations on a given trial (in line with rapid and parallel processing characteristics of object-based attention; de-Wit *et al,* 2011).

Awareness was then assessed by first revealing to the participants the nature of the objects and requiring them to view the stimuli from the attention task a second time. In this second phase, participants made a decision on each trial as to whether the cue and target appeared within a single object or in different objects in a confidence-rating signal-detection procedure.

6.10.1. Methods

Participants

Twenty naive observers participated, and all gave their informed consent. Participants were undergraduate and postgraduate students recruited through the Durham University Psychology Department's participant pool scheme, and were awarded either course credits for their participation or a small financial compensation.

Materials

See previous experiment.

Procedure

See previous experiment. There were only four potential object locations: above, below, left or right of fixation, one placeholder was located at either end of both figures and the spatial distance between these placeholders was equated. For each block of trials, the objects were presented an equal number of times vertically and horizontally and the order of this was randomised within each block. On each trial, the location of the objects was determined randomly with equal probability. See **figure 6.6** (**a** and **b**) for a depiction of the display sequence and a simulation of the observers' perception during the sequence, respectively. **Figure 6.6c** also shows examples of the different positions and orientations which the objects could assume.

Participants completed 10 practice trials followed by three blocks of 120 experimental trials. Participants were then asked an open question to probe their visual experience of the stimuli: they were asked to describe anything they noticed about the flickering background on which the white flash (cue) and coloured disc (target) were presented. After answering the question, participants were then shown the display with a much reduced alternation rate of 4 Hz which explicitly revealed the nature of the objects in the display.

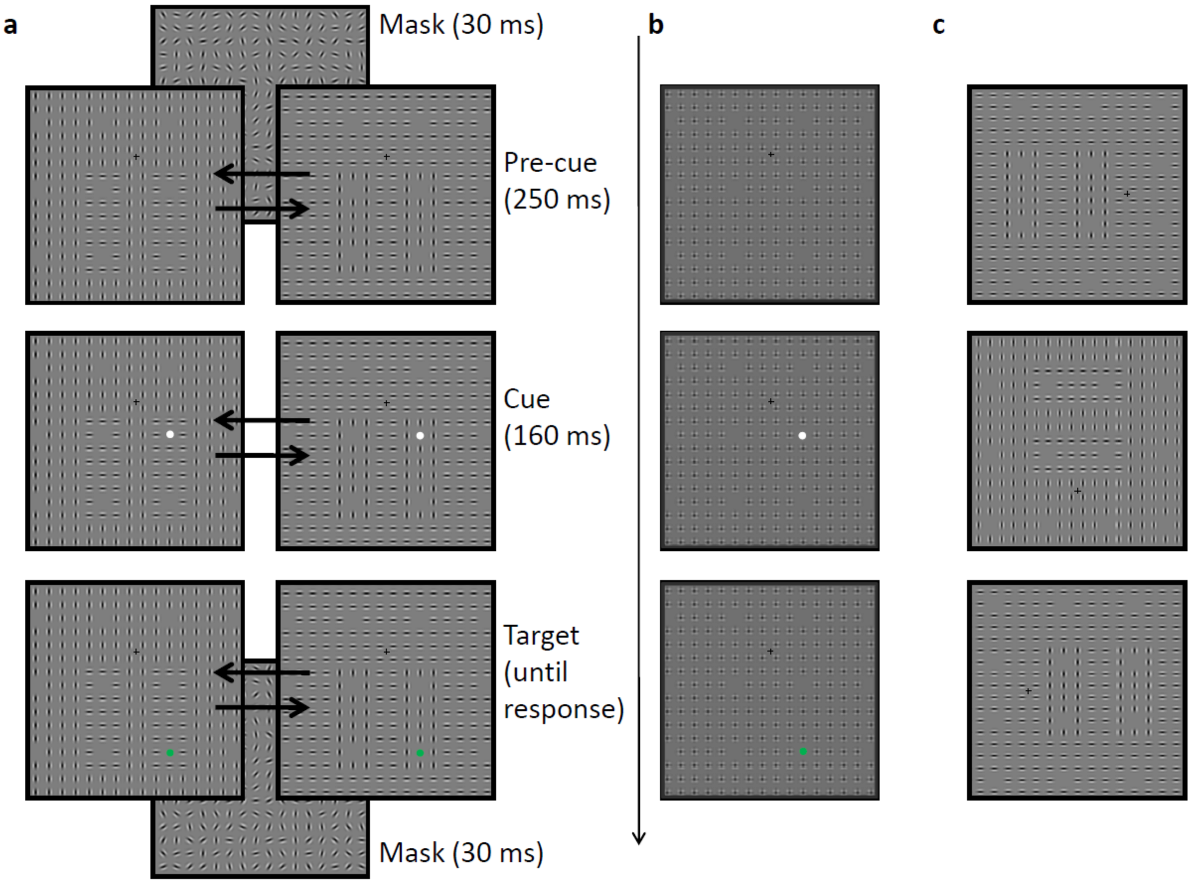

Figure 6.6 Illustrations of the stimuli and the temporal sequence of experiment 5; a simulated observer's perception of those events; and three examples of the different object positions and orientations. **a)** Full illustration of the trial sequence. Note that each box only shows a magnified portion of the entire stimulus display, focussed on the objects. Here the objects appear below the fixation cross in a vertically-aligned arrangement and the target is invalidly cued within those objects. Double arrows indicate that the two frames are presented continually in alternation at a frequency of 16.7 Hz. In the signal-detection task, the final two frames (target frames) were only presented for a limited amount of time that was calibrated for each participant (see ***Methods***) **b)** A simulated observers perception of the sequence shown in **a**. Participants are not aware of the presence of the figures. **c)** Examples of the three object positions not shown in **a;** relative to the fixation cross, they are left, above and right. The cue and target on each trial would only be presented within the 4 placeholders associated with the figures.

The second phase of the experiment determined whether participants could identify "within-object" trials from "between-object" trials in a confidence-rating signal-detection procedure. Participants were presented with an additional 3 blocks of 120 trials, preceded by 10 practice trials, each containing the same proportion of valid, invalid-within and invalid-between trials, and the same number of horizontally and vertically presented objects as in the previous attention task. Other randomly-determined parameters (e.g. object position and orientation) and temporal characteristics remained consistent with the attention task, except that following the onset of the target, the display would only remain on the screen for a limited amount of time. This was automatically determined individually for each participant by obtaining their largest RT from the attention task following the removal of outliers (removal criteria described in full in the **Results** section). Participants had to indicate on which trials the cue and target appeared in the same object or in different objects by pressing one of two keys, and then to rate their confidence with that judgment on a scale of 1-4, by pressing one of four keys.

A previous experiment, experiment 4 here, the essence of which was the same as the present study, strongly suggested that objects concealed from awareness in the manner described above could mediate object-based attention without awareness. In the previous experiment participants were simply required to report whether or not the flickering display contained objects defined by an area of non-uniformity in the signal detection task. We did not reveal the objects to the participants prior to this task (i.e. they were not shown the display at a rate of 4Hz), to minimise any discrepancy in their knowledge between the attention task and signal detection task, as such factors may lead to significant overestimations of sensitivity in the signal-detection task (Vermeiren & Cleeremans, 2012). In the present study, however, we aimed to adopt a more stringent assessment of awareness, by revealing the exact nature of the objects to the participants prior to the signal-detection task, notwithstanding the likelihood that this manipulation would result in an overestimation of sensitivity relative to the attention task.

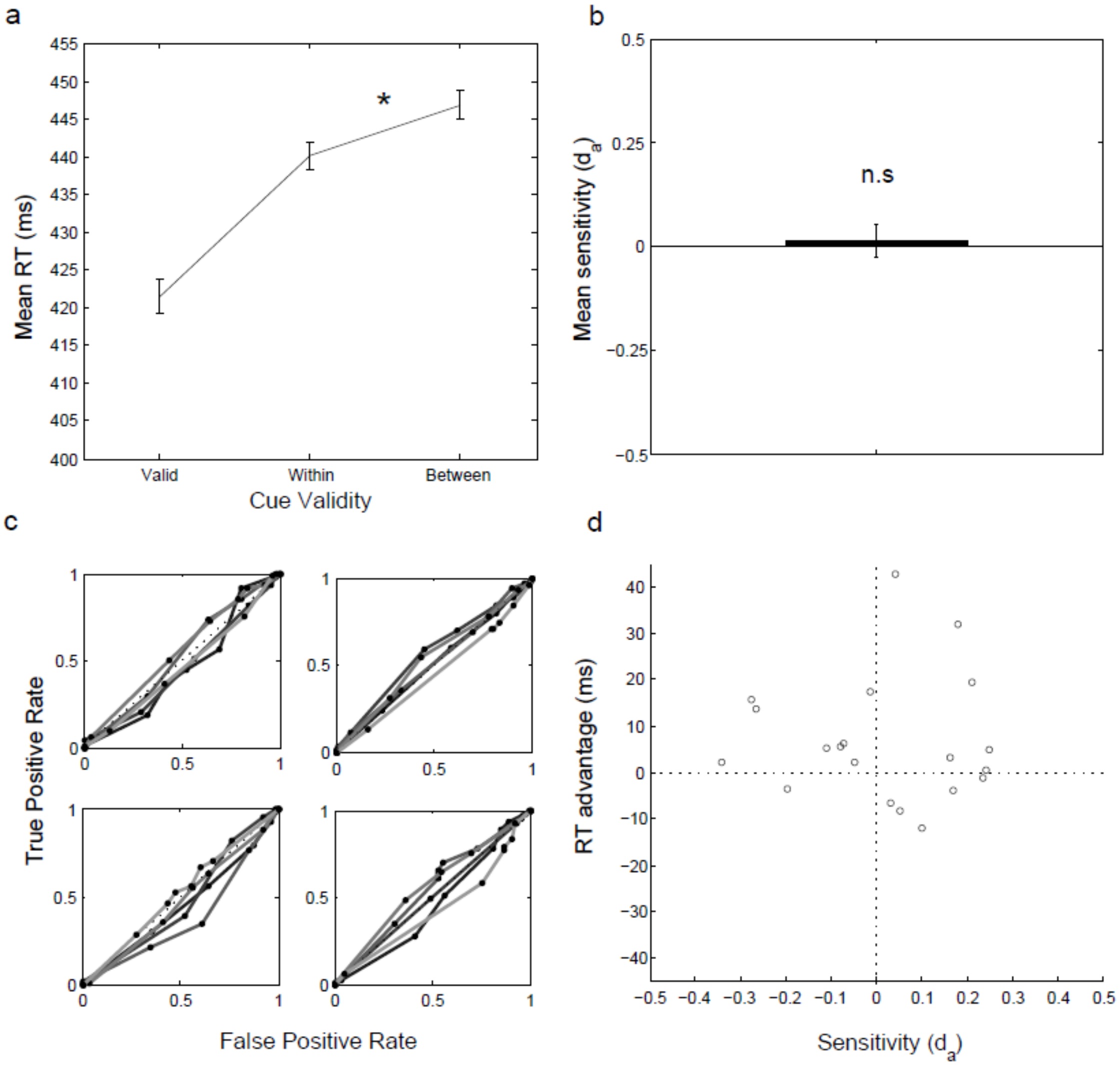

Figure 6.7: Results from both the attention task and signal-detection task from experiment 5. **a)** RT results from the attention task (N=20). Participants were quicker to identify the colour of a target when it appeared within the same object as a preceding cue compared with when it appeared on a different object, whilst spatial distance was equated. This is true despite participants denying any awareness of those objects. Asterisk denotes a significant paired t-test at the 5% level. Error bars show 1 standard error of the mean with between-participant variance omitted **b)** Average sensitivity (d_a) to the difference between within- and between-object trials. This value is not statistically different from 0, indicating absence of awareness. **c)** Individual participants' ROC. All participants show an approximate linear plot indicating that they could not successfully maximise hit rate whilst minimising false alarm rate. Each set of axes represents 5 of the total 20 participants' ROC curves. **d)** Scatterplot of each participant's RT advantage (between-object RT - within-object RT) versus sensitivity. There is no observable association between the participants' ability to discriminate within-object trials from between-object trials and their difference in RT between the two conditions.

6.10.2. Results and discussion

RTs were trimmed using the same methods as in the previous experiments. Overall, 8.6% of all trials were discarded using these methods. A within-participant ANOVA with the single factor Cue Validity was conducted on the mean values of the remaining RTs, with overall means of 421.5ms (valid), 440.1ms (invalid-within) and 446.7ms (invalid-between) as shown in **figure 6.7a**. The main effect was significant ($F_{(2,38)}$=28.99, p<0.001) indicating that the cue had a different effect on participants' RTs depending on its position relative to the target and the figures. In the key analysis, a paired t-test revealed that RTs were shorter on invalid-within than for invalid-between trials ($t_{(19)}$= 2.24, p=0.037), indicating that participants were faster to respond to targets that appeared within the same object as the preceding cue relative to those that appeared on a different object. No significant effect of accuracy (valid: 96.3%, invalid-within: 96.1%, and invalid-between: 96.8%) was found between the conditions of cue validity ($F_{(2,38)}$=0.45, p=0.639), indicating no trade-off between RT and accuracy.

In response to the open question asked regarding the content of the flickering background, most participants described it as being composed of "flickering crosses" or "flickering lines" and some would remark that there were parts of the background that were "missing" (the placeholders). None, however, made any comments that could be interpreted in any way as awareness of the objects. Data from the subsequent signal-detection task were used to formally measure participants' sensitivity to the objects. The task measured participants' ability to distinguish two categories of trial ("within-object" and "between-object") that each occurred an equal number of times. Participants indicated which category they judged each trial to belong to by pressing one of two keys, and then rated their confidence with that judgment on a scale of 1-4. For the purposes of signal-detection, this is analogous to a yes/no design, in which "between-object" trials are treated as "signal-present" and "within-object" trials are treated as "signal-absent (noise)". Participants' responses were first assigned to one of two categories (within/between) and the additional confidence report by the participant on this decision (on an integer scale of 1-4) thus provided a total scale of 8 responses: ranging from very confident "within-object" (1) to very confident "between-object" (8). Sensitivity was calculated by tabulating the number of responses for each of these eight confidence levels for both "within-object" and "between-object" trials. The discriminability index d_a was calculated from these data using the software RScorePlus (Harvey, 2002) to fit a Gaussian unequal variance model.

d_a allows for unequal variance and is equivalent to d' in the case of equal variance. A higher d_a indicates a greater sensitivity to the signal, and a d_a of zero indicates no sensitivity. A negative d_a can represent an ability of an observer to discriminate the two conditions, but the conditions are labelled incorrectly by the observer (e.g. in the case of this experiment, responding consistently with the "within-object" response for "between-object" trials, and *vice versa*), however it can also simply be (and is more likely to be in this case) a consequence of sampling error.

Overall, participants' d_a (shown in **figure 6.7b**) did not differ significantly from zero (mean d_a = 0.01; $t_{(19)}$ = 0.32, p = 0.75), indicating that participants could not discriminate the two conditions of within-object and between-object trials, and hence it is extremely unlikely that they had any awareness of the objects. Receiver operating characteristics (ROCs) were also computed from the same data for each participant. Each curve contains 7 points (as a scale of *n* criteria, in this case 8, determines *n-1* points on the curve), with each representing a single criterion that distinguishes one rating from the immediately lower rating (e.g. rating 4 from rating 3, or rating 8 from rating 7). The ROC curves are plotted with true positive rate (hit rate) vs. false positive rate (false alarm rate) in **figure 6.7c**; the more linear the plot, the less able the observer is to differentiate the two conditions. As shown in **figure 6.7c**, no participant showed any ability to maximise hit rate whilst minimising false alarm rate (as would be indicated by a bowed curve), indicating that they could not accurately distinguish the conditions that were driving the object-based attention effects in the previous task.

Figure 6.7d also shows a scatterplot of points for each individual participant's within-object RT advantage (calculated from between-object RT – within-object RT) versus sensitivity (d_a). A parametric correlation test between these two variables is not significant (r(18)=-0.019, p=0.935), clearly indicating no association between the awareness of the objects and their effect on attention.

6.11. *Experiment 6 – adopting the most stringent assessment of awareness*

Awareness in the previous task was assessed by measuring participants' discrimination of "within-object" trials from "between-object" trials. Although with this setup there is a great amount of parity between the attention task and the signal detection task (an important consideration when assessing effects of cues/primes without awareness; Reingold & Merikle, 1988), the dissociation in the results may have stemmed from an inappropriate level of task difficulty in the signal detection experiment as participants needed to encode and combine separate information regarding the objects, cue and target in order to make a successful response. It is important, therefore, to demonstrate an absence of awareness of the objects when participants do not have such a cognitive demand. Awareness was thus assessed in the following experiment by first revealing to the participants the nature of the objects and requiring them to view the stimuli from the attention task a second time. In this second phase, however, only half of the trials contained the objects and it was the participants' task to distinguish these "object-present" trials from those in which the objects were absent in a confidence-rating signal-detection procedure. Although this does not have parity between the two tasks, it is arguably the most stringent assessment of the absence of awareness that is possible in this experiment. The following experiment is published in *Psychological Science*. The citation is:

Norman, L. J., Heywood, C. A., & Kentridge, R. W. (2013). Object-based attention without awareness. *Psychological Science, 24,* 836-843

It is reproduced here in its entirety as part of this chapter. As the text has not been altered, some information will have already been discussed in the preceding sections of this chapter.

6.11.1. Abstract

Attention and awareness are often considered to be related. Some forms of attention can, however, facilitate the processing of stimuli that, nevertheless, remain unseen. It is unclear whether this dissociation extends beyond selection on the basis of primitive properties such as spatial location to situations in which there are more complex bases for attentional selection. The experiment described here shows that attentional selection at the level of objects can take place without giving rise to awareness of those objects. Pairs of objects were continually masked, rendering them invisible to participants

performing a cued-target discrimination task. When the cue and target occurred within the same object, discrimination was faster than when they occurred in different objects at the same spatial separation. Participants reported no awareness of the objects and were unable to detect them in a signal-detection task. Object-based attention, therefore, is not sufficient for object-awareness.

6.11.2. Introduction

There is an association between the phenomena of consciousness and attention so irresistible that one could readily conclude they were inextricably part of the same process (James, 1980). As Lamme (2003) notes, both are selective: not all visual input reaches awareness, and only a fraction of it is treated with the efficacy that is offered by selective attention. Thus it had long been assumed that prioritization of information by attention was both necessary and sufficient for consciousness (Mole, 2008). Remarkable demonstrations of inattentional blindness, in which otherwise conspicuous visual events are rendered invisible with diverted attention, bolstered this assumption (Mack & Rock, 1998).

Visual spatial attention reflects the voluntary or involuntary prioritization of information in a selected part of a visual scene (Posner, 1980). Experimentally, Posner's cueing task, in which a cue facilitates performance by speeding the discrimination of a target in the same location, has been considered the benchmark measurement of covert visual attention. It has now been shown, however, that attention in this manner is not *sufficient* for generating visual awareness. Kentridge, Heywood & Weiskrantz (1999) demonstrated this in a blindsight patient who declares no awareness of visual experience in his right hemifield as a result of unilateral striate cortex damage and yet retains remarkably accurate performance in some forced-choice discriminations made within that part of his visual field (Weiskrantz, 1986). Selective attentional modulation was observed in this patient in his responses to *cued* stimuli compared with *uncued* stimuli in his blind field, in very much the same way as a normal "aware" observer in a Posner (1980) task. Similar effects have since been found in normal observers, when a masked unseen prime has a greater effect on the subsequent discrimination of a target when attention is directed towards it (Kentridge, Nijboer & Heywood, 2008; see also Sumner, Tsai, Yu & Nachev, 2006). This experimental evidence strongly suggests that selective attention is not sufficient to give rise to awareness. In parallel, magnetoencephalographic recordings have also pointed to independent neural mechanisms regulating spatial attention and awareness in normal observers (Wyart & Tallon-Baudry, 2008).

In the spotlight model of attention (Posner, 1980), selection is based on a simple spatial primitive in which attention is focussed on a single point in space and spreads uniformly around it. Attention is, however, not limited to such simple, purely spatial, primitives; objects of arbitrary shape can form the "units" of attentional selection (Duncan, 1984; Egly, Driver & Rafal, 1994). Egly and colleagues (1994) demonstrated the importance of objects for the deployment of attention in the classic modification of Posner's (1980) task, in which visual discriminations were shown to be more rapid when the target was seen to be within the same object as the preceding cue compared with when it was seen to be within a separate object, despite both cue-target pairings being equidistant. This can be explained by a model in which attentional selection operates on the elementary figures that are preattentively segmented by the visual system[10]. It has been argued that in the cases of dissociation of attention from awareness, it is only *spatial* attention that has been manipulated, whereas awareness has been assessed typically on the basis of the visibility of *objects* (Mole, 2008). In other words, the unit of selection and the object of awareness may not have been truly equivalent in studies claiming to demonstrate dissociations between attention and awareness.

The motivation of the present study was to determine whether objects can act as units of attentional selection even when they are not consciously seen. This would be striking, not only because object-based attention involves a level of sophistication beyond simple spatial selection, but also because there is clear parity between the objects of attention and awareness in the present experiment that may be lacking in tasks solely employing simple spatial selection. In the experiment described here, objects were defined by an orientation contrast to their background, but, crucially, the orientations of the texture elements were continually reversed. Orientation-reversing stimuli of this type have previously revealed that the perception of a border between two adjacent regions of

---

[10] The authors do not claim to disentangle a purely object-based form of attention from one that involves the selective spreading of what is fundamentally spatial attention within an object (see Martinez *et al,* 2006). The only aim of this study was to show that the process of segregating visual information into objects for the purposes of attention was not a sufficient precondition for the awareness of those objects.

texture persists despite the two regions being continually masked (Norman, Heywood & Kentridge, 2011). In the present experiment the orientations were reversed at a frequency above which the conscious perception of the contour also vanishes. We hypothesised that the contour between the objects and background may nonetheless be processed at a level that allows object-based attention, whereas any awareness of those objects would be prevented by the continual masking of the stimuli.

Participants completed a standard Egly *et al* (1994) cueing task, in which they discriminated the colour of a target that was validly cued (50% of trials), invalidly cued within the same object as the cue (25%) or invalidly cued in another object (25%). The objects appeared at unpredictable locations and assumed unpredictable orientations on a given trial (in line with rapid and parallel processing characteristics of object-based attention; de-Wit, Cole, Kentridge & Milner, 2011).

Awareness was then assessed by first revealing to the participants the nature of the objects and requiring them to view the stimuli from the attention task a second time. In this second phase, however, only half of the trials contained the objects and it was the participants' task to distinguish these "object-present" trials from those in which the objects were absent in a confidence-rating signal-detection procedure.

6.11.3. Methods

Participants

Ten naive observers participated, and all gave their informed consent. Participants were undergraduate and postgraduate students recruited through the Durham University Psychology Department's participant pool scheme, and were awarded either course credits for their participation or a small financial compensation.

Materials

Stimuli in all experiments were generated using a Cambridge Research Systems ViSaGe Graphics System and were presented on a gamma-corrected ViewSonic 17'' display monitor viewed at a distance of 41cm (participants rested their head on a chin rest). The background had a luminance of 50cdm^{-2}. The screen resolution was set to 1024 x 768 pixels with a refresh rate of 100 Hz. The ViSaGe Graphics System ensured that stimulus display and response timing were time-locked with the monitor's refresh rate.

Procedure

Participants fixated a central cross. Following a warning tone (frequency = 800Hz, duration = 100ms) and a delay of 500ms, a lattice (17.2° in width and in height) centred on the fixation cross, consisting of 22 x 22 Gabor patches arranged in a uniform array, appeared. Each Gabor patch had a diameter of 0.4°, a spatial frequency of 3.75 cycles/degree and an envelope with a standard deviation of 0.2°. Each patch had a Michelson contrast of 90% and was separated from its neighbours by 0.4°. See **figure 6.8a** for illustrations of the stimuli and procedure. For 30ms, these Gabor patches were presented, each with a randomly determined orientation, as a mask before the onset of the objects. Immediately following this, the patches would continually alternate between vertical and horizontal orientations at 16.7Hz. A pair of identical rectangular objects (composed of 8 x 3 Gabor patches and presented either above, below, to the left or to the right of fixation at a distance of one Gabor patch spacing) was formed by an orientation contrast of 90° to the background. Thus, the objects were always defined by an orientation contrast of 90° to their background even as the orientations of *all* the Gabor patches in the display alternated. The flicker rate of 16.7Hz implies that each frame is present on screen for 30ms; a short duration but one which is quite typical of stimuli in forward and backward masking paradigms. The effectiveness of this flicker rate in concealing the objects from awareness was also verified in a number of pilot studies.

In 16 of the Gabor positions, located in a 4 x 4 arrangement (i.e. at every 4th position, vertically and horizontally) the Gabor patches were absent, serving as placeholders for cues or targets. This ensured that for each of the 4 possible object locations (above, below, left or right of fixation), one placeholder was located at either end of both figures and the spatial distance between these placeholders was equated. For each block of trials, the objects were presented an equal number of times vertically and horizontally and the order of this was randomised within each block. On each trial, the location of the objects was determined randomly with equal probability.

250ms after the object onset, a cue (a white disc: luminance = 158 cdm^{-2}, 0.4° in diameter) appeared for 160ms in one of the four placeholders associated with the object positions (determined randomly with equal probability on each trial). Following the offset of the cue, the target disc (0.4° in diameter) appeared in one of three locations, and would either be red (CIE 1931 x, y coordinates of 0.40, 0.31 with a luminance of 72 cdm^{-2}) or green

(CIE 1931 x, y coordinates of 0.30, 0.59 with a luminance of 81 cdm^{-2}). In valid trials (50% of all trials) the target would appear in the same position as the cue. In invalid-within trials (25%) the target appeared in the adjacent placeholder within the same object in which the cue had been presented. In invalid-between trials (25%) the target appeared in the adjacent placeholder that was within a different object. The colour of the target was determined randomly with equal probability on each trial. Participants were instructed to indicate the colour of the target disc (red/green) by pressing one of two buttons. The target remained on the screen until a response was made; following which the noise mask of random orientations was presented again for a further 30ms, ending the trial. See **figure 6.8** (**a** and **b**) for a depiction of the display sequence and a simulation of the observers' perception during the sequence, respectively. **Figure 6.8c** also shows examples of the different positions and orientations which the objects could assume.

Participants completed 10 practice trials followed by three blocks of 128 experimental trials. Participants were then asked an open question to probe their visual experience of the stimuli: they were asked to describe anything they noticed about the flickering background on which the white flash (cue) and coloured disc (target) were presented. After answering the question, participants were then shown the display with a much reduced alternation rate of 4 Hz which explicitly revealed the nature of the objects in the display.

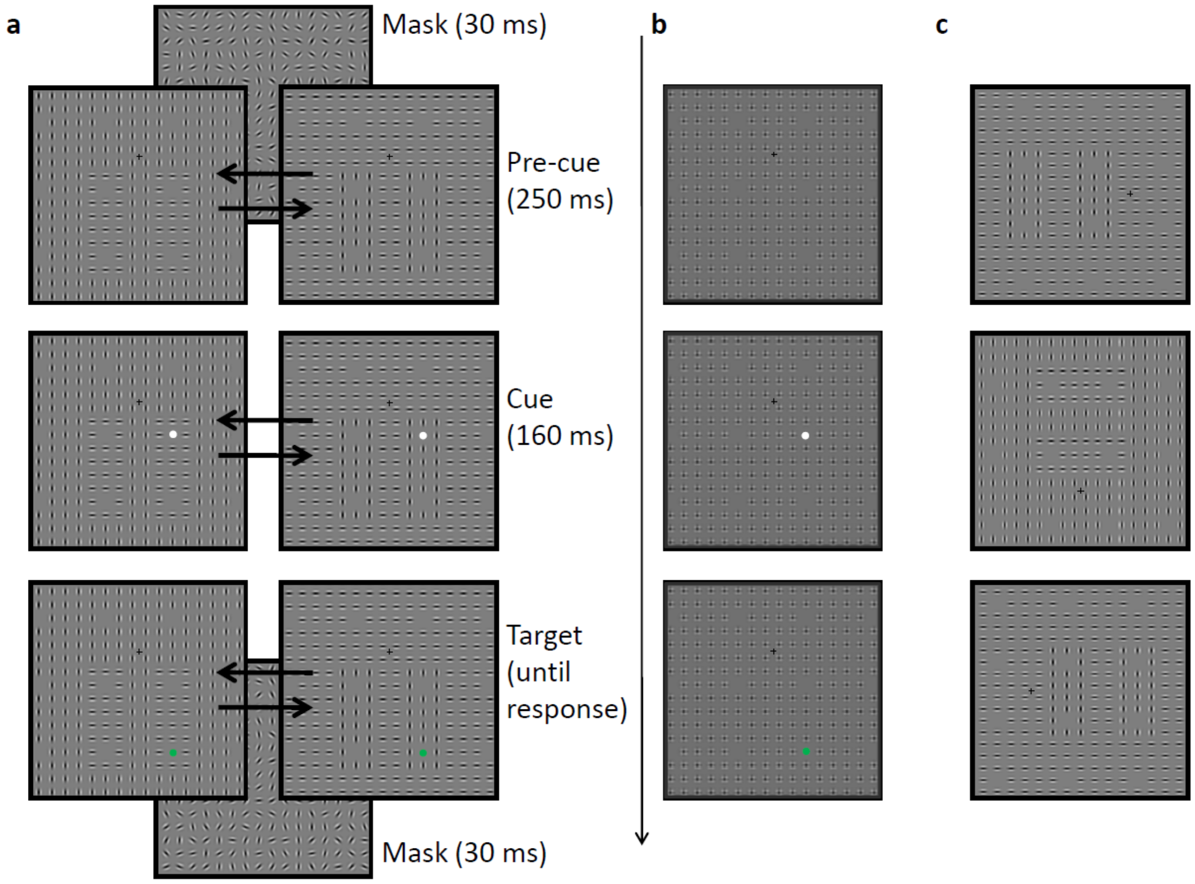

Figure 6.8: Illustrations of the stimuli and the temporal sequence of experiment 6; a simulated observer's perception of those events; and three examples of the different object positions and orientations. **a)** Full illustration of the trial sequence. Note that each box only shows a magnified portion of the entire stimulus display, focussed on the objects. Here the objects appear below the fixation cross in a vertically-aligned arrangement and the target is invalidly cued within those objects. Double arrows indicate that the two frames are presented continually in alternation at a frequency of 16.7Hz. In the signal-detection task, the final two frames (target frames) were only presented for a limited amount of time that was calibrated for each participant (see ***Methods***) **b)** A simulated observers perception of the sequence shown in **a**. Participants are not aware of the presence of the figures. **c)** Examples of the three object positions not shown in **a;** relative to the fixation cross, they are left, above and right. The cue and target on each trial would only be presented within the 4 placeholders associated with the figures.

The second phase of the experiment determined whether participants were sensitive to the presence of the objects in a confidence-rating signal-detection procedure. Participants were presented with an additional 3 blocks of 128 trials, preceded by 10 practice trials, in which the objects were present only in half the number of trials. When the objects were absent, all orientations in the display were homogenous and alternated every 30ms between horizontal and vertical. All randomly-determined parameters (e.g. object position and orientation, when present) and temporal characteristics remained consistent with the attention task, except that following the onset of the target, the display would only remain on the screen for a limited amount of time. This was automatically determined individually for each participant by obtaining their largest RT from the attention task following the removal of outliers (removal criteria described in full in the **Results** section). The mean of these values was 937ms (SD=158, max=1280ms, min=739ms). Participants had to indicate on which trials the objects were present by pressing one of two keys, and then to rate their confidence with that judgment on a scale of 1-4, by pressing one of four keys.

6.11.4. Results

In the attention task only correct trials were analysed. The RTs were trimmed by first removing those that exceeded 1500ms, interpreted as unsuccessful button-presses or momentary lapses in concentration, or were less than 200ms, interpreted as anticipatory responses. The remaining data that fell outside 2 standard deviations from the mean per condition per participant were removed as outliers. Overall, 6.2% of all trials were discarded. A within-participant ANOVA with the single factor Cue Validity was conducted on the mean values of the remaining RTs, with overall means of 456.7ms (valid), 482.9ms (invalid-within) and 488.8ms (invalid-between) as shown in **figure 6.9a**. The main effect was significant ($F_{(1.11,9.95)}=25.61$, $p<0.001$; Greenhouse-Geisser corrected) indicating that the cue had a different effect on participants' RTs depending on its position relative to the target and the figures. In the key analysis, a paired t-test revealed that RTs were shorter on invalid-within than for invalid-between trials ($t_{(9)}= 3.41$, $p=0.008$), indicating that participants were faster to respond to targets that appeared within the same object as the preceding cue relative to those that appeared on a different object. No significant effect of accuracy (valid: 97.7%, invalid-within: 97.5%, and invalid-between: 98.2%) was found between the conditions of cue validity ($F_{(2,18)}=0.48$, $p=0.626$), indicating no trade-off between RT and accuracy.

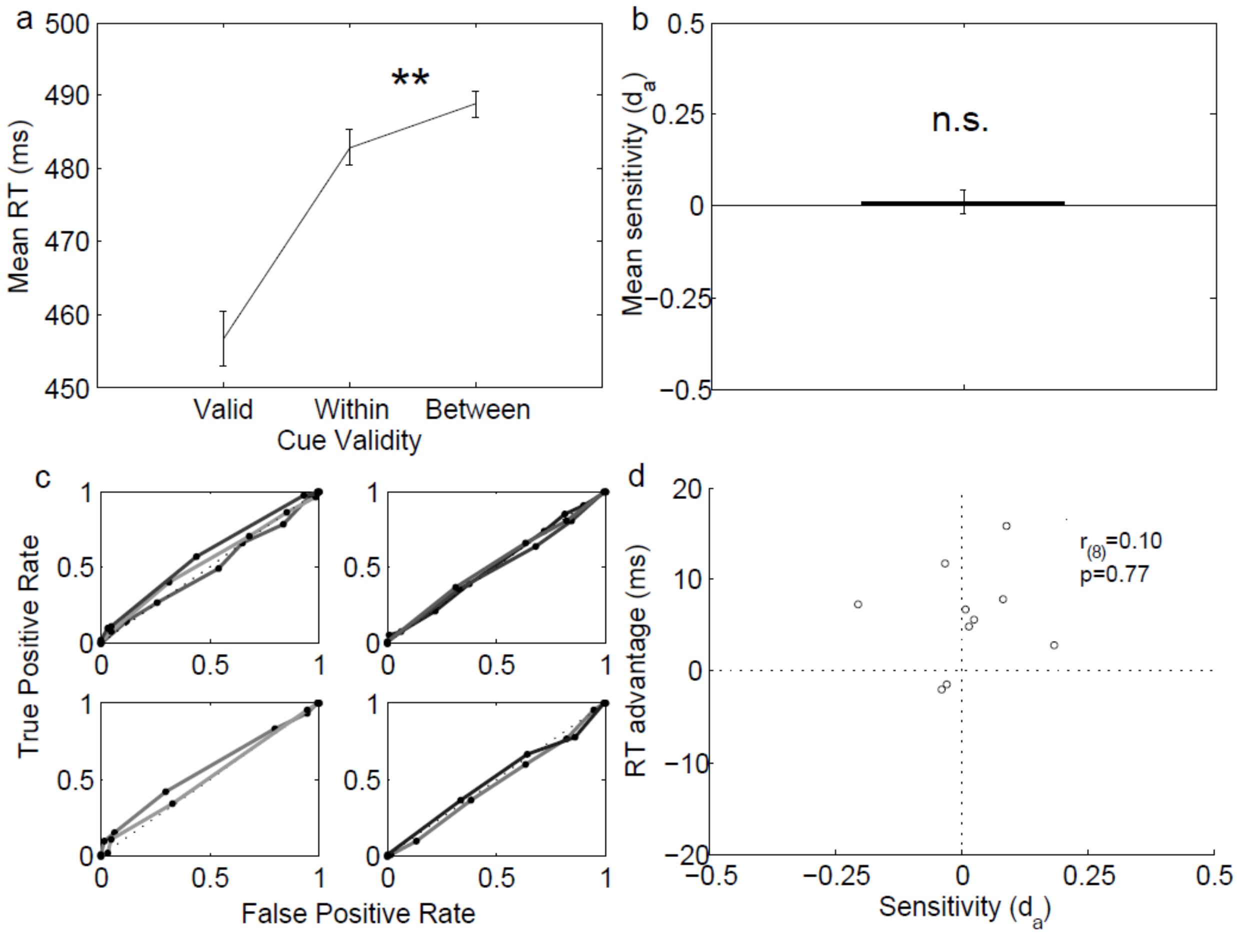

Figure 6.9: Results from both the attention task and signal-detection task from experiment 6. **a)** RT results from the attention task (N= 10). Participants were quicker to identify the colour of a target when it appeared within the same object as a preceding cue compared with when it appeared on a different object, whilst spatial distance was equated. This is true despite participants denying any awareness of those objects. A double asterisk denotes a significant paired t-test at the 1% level. Error bars show 1 standard error of the mean with between-participant variance omitted **b)** Average sensitivity (d_a) to the preence of the objects (vs. their absence) in the signal detection task. This value is not statistically different from 0, indicating absence of awareness. **c)** Individual participants' ROC. All participants show an approximate linear plot indicating that they could not successfully maximise hit rate whilst minimising false alarm rate. Each set of axes represents either 3 or 2 of the total 10 participants' ROC curves. **d)** Scatterplot of each participant's RT advantage (between-object RT - within-object RT) versus sensitivity. There is no association between the participants' ability to detect the presence of the objects and their difference in RT between the within-object and between-object conditions.

In response to the open question asked regarding the content of the flickering background, most participants described it as being composed of “flickering crosses” or “flickering lines” and some would remark that there were parts of the background that were “missing” (the placeholders). None, however, made any comments that could be interpreted in any way as awareness of the objects. Data from the subsequent signal-detection task were used to formally measure participants’ sensitivity to the objects. The task measured participants’ ability to distinguish two categories of trial (“objects-present” and “objects-absent”) that each occurred an equal number of times. Participants indicated which category they judged each trial to belong to by pressing one of two keys, and then rated their confidence with that judgment on a scale of 1-4. Participants’ responses were first assigned to one of two categories (present/absent) and the additional confidence report by the participant on this decision (on an integer scale of 1-4) thus provided a total scale of 8 responses: ranging from very confident “objects-present” (1) to very confident “objects-absent” (8). Sensitivity was calculated by tabulating the number of responses for each of these eight confidence levels for both “objects-present” and “objects-absent” trials. The discriminability index d_a was calculated from these data using the software RScorePlus (Harvey, 2002) to fit a Gaussian unequal variance model. d_a allows for unequal variance and is equivalent to d' in the case of equal variance. A higher d_a indicates a greater sensitivity to the signal, and a d_a of zero indicates no sensitivity. A negative d_a can represent an ability of an observer to discriminate the two conditions, but the conditions are labelled incorrectly by the observer (e.g. in the case of this experiment, responding consistently with the “objects-present” response for “objects-absent” trials, and *vice versa*), however it can also simply be (and is more likely to be in this case) a consequence of sampling error.

Overall, participants’ d_a (shown in **figure 6.9b**) did not differ significantly from zero (mean $d_a = 0.01$; $t_{(9)} = 0.32$, $p = 0.75$), indicating that participants could not discriminate those trials in which the objects were present from those in which they were absent, and hence it is extremely unlikely that they had any awareness of the objects. Receiver operating characteristics (ROCs) were also computed from the same data for each participant. Each curve contains 7 points (as a scale of *n* criteria, in this case 8, determines *n-1* points on the curve), with each representing a single criterion that distinguishes one rating from the immediately lower rating (e.g. rating 4 from rating 3, or rating 8 from rating 7). The ROC curves are plotted with true positive rate (hit rate) vs. false positive rate (false alarm rate) in **figure 6.9c**; the more linear the plot, the less able the observer is to differentiate the two conditions. As shown in **figure 6.9c**, no participant showed any ability to maximise hit rate whilst minimising false alarm rate (as would be indicated by a bowed curve), indicating that they could not accurately distinguish the conditions that were driving the object-based attention effects in the previous task.

Figure 6.9d also shows a scatterplot of points for each individual participant's within-object RT advantage (calculated from between-object RT – within-object RT) versus sensitivity (d_a). A parametric correlation test between these two variables is not significant ($r_{(8)}$=0.10, p=0.78), clearly indicating no association between the awareness of the objects and their effect on attention.

6.11.5. Discussion

In this experiment, targets appearing within the same object as a cue were processed more rapidly than those appearing on a different object. This is a standard demonstration of object-based attention (Egly *et al,* 1994), although participants showed no evidence of any conscious access to the objects as revealed through a signal-detection task. This is in line with the more general notion that engaging attention is not sufficient for awareness, which has previously been demonstrated (Kentridge *et al,* 1999; 2004; 2008; Sumner *et al,* 2006; Kanai, Tsuchiya & Verstraten, 2006; Koch & Tsuchiya, 2007) and that attention and awareness have distinct neural signatures (Wyart & Tallon-Baudry, 2008). The magnitudes of the within-object advantages reported here are small but by no means atypical of those found using this paradigm with visible objects (see Reppa, Schmidt & Leek, 2012). Crucially, the present experiments refute the potential claims that previous work had only manipulated an early purely spatial form of attention, in which attention is not directed at an object *per se*, but rather simply the space which it occupies (Mole, 2008). Additionally, it is very important to note that the participants' inability to detect the objects above-chance could not simply be attributed to a memory failure, as the participants were at liberty to make responses in the signal detection task at the instant they became aware of the objects.

There are many issues to consider when choosing the most appropriate way to assess awareness in experiments that exploring unconscious attentional effects (Vermeiren & Cleeremans, 2012). In the present experiment we report the most straightforward measure – a test of participants' ability to discriminate objects' presence versus their absence. The critical property that determines the unconscious attention effect, however, is the objects' spatial location and orientation together with the relative positions of the cue and target on each trial. Even if participants are unable to discriminate the presence and absence of objects per se they might conceivably retain a conscious impression as to whether cues and targets appeared within a single object. It could be argued, therefore, that in the signal detection task, participants should in fact be required to discriminate "within-object" from "between-object" trials, on the basis that this maximises the parity between the tasks measuring attention and awareness (Reingold & Merikle, 1988). We conducted a separate experiment using the same general methods with an independent sample of 20 participants with this alternate signal detection task. The results are reassuring, as the participants could not discriminate the two types of trial in a signal detection task and again demonstrated a reliable within-object reaction time advantage[11]. Thus, we have demonstrated using two variations of a signal detection task that both the objects' presence and their spatial relationship with cues and targets are concealed from awareness. The signal detection task used in the present study is arguably the most stringent assessment of awareness, as the task has a relatively low cognitive demand; participants do not need to encode and combine separate information regarding the objects, cue and target in order to make a successful response, as the task requires only knowledge concerning the objects' presence.

An important aspect of normal object-based attention is that it is effortless, and is deployed rapidly across a visual scene (de-Wit *et al,* 2011), which must certainly be true if we are to assume that these effects have any bearing on the mechanisms of everyday visual perception. It appears to be an *automatic* rather than a voluntarily controlled process of selection. In the present experiment objects could appear randomly in one of 4 locations with vertical or horizontal orientation. Observing an object-based effect under these conditions is in keeping with the current understanding the operation of automatic preattentive scene segmentation for the purposes of object-based attention. If this finding does reflect an unconscious deployment of exogenous object-based attention, it also has wider implications for our understanding of the relationship between attention and

---

[11] The result of the critical paired t-test between the within- and between-object trials was $t_{(19)}= 2.24$, $p=0.037$.

awareness; Chica and colleagues (2011) have shown that behavioural and electrophysiological signatures of attention and awareness dissociate when attention is voluntarily controlled but show stronger correlation when attention is under exogenous control, lending weight to the suggestion that awareness automatically follows exogenous attention. The results of our experiment suggest that although exogenous attention may be necessary for awareness – exogenously controlled attention can act in the absence of awareness.

The object-based attention effects observed in the present study stem from simple segmentation processes; however, the visual system also uses grouping principles to infer the true nature of the environment when complete segmentation information is not immediately available (i.e. when objects may be partially occluded). Consequently, object-based attention effects are still found for partially-occluded objects (Moore, Yantis & Vaughan, 1998). An important step in future research, therefore, will be to determine whether such grouping principles that extend beyond simple segmentation processes impose a limitation on the functionality of object-based attention in the absence of awareness.

The visual system is purported to process information of an object's structure for purposes that do not automatically result in awareness. This has been illustrated in cases of visual form agnosia, in which patients have no conscious access to the shapes of objects, arising from bilateral damage to the lateral occipital complex (LOC), an area selective to object shape (Grill-Spector, Kushnir, Edelman, Itzchak & Malach, 1998). Yet, paradoxically, these patients may retain an ability to manipulate those objects appropriately in accordance with their shape (Milner *et al,* 1991). Such a dissociation is believed to reflect the division of labour between the dorsal (subserving unconscious guiding of action towards objects) and ventral (subserving conscious perception and recognition of objects) streams of visual processing (Milner & Goodale, 1995). de-Wit, Kentridge and Milner (2009a) showed that object-based attention could not be engaged in a patient with bilateral LOC damage despite their having an otherwise functional attention system. This highlights the necessity of area LOC, a region of the ventral stream, in representing form for the purposes of object-based attention. Although the ventral stream is viewed as predominantly a *conscious* processing stream, there are occasions in which activity within it correlates with stimulus information despite an absence of awareness of the stimuli (Dehaene *et al,* 2001; Sterzer, Haynes &

Rees, 2008). Indeed, Sterzer and colleagues' (2008) discovered ventral activity that differentiated images of faces from houses even when the stimuli were not entering awareness. This high-level but unconscious categorization by the visual system is arguably a more complex process than that which determines the orientation of a pair of rectangles in an object-based attention task, and so it is not surprising to find evidence of unconscious ventral stream operations in the present study. What is surprising, however, is that this unconscious ventral stream activity is capable of directing attention. One possible explanation is that the forward and backward masking employed in this study substantially reduces the object-related activity in ventral areas to a level below the threshold of visual awareness but not that of attention. Alternatively, however, the feedback from these ventral areas to primary visual cortex may be of critical importance (e.g. see Fahrenfort, Scholte & Lamme, 2007), something which is disrupted by the continual masking in the current experiment but which is, perhaps, not necessary for the operation of object-based attention.

6.12. General discussion of study five

The experiments reported here have shown numerous instances of object-based attention without awareness of the objects. In the initial two experiments, the assessment of awareness came only from the participants' report of their subjective experience of the experiment, which was acquired through a series of post-experiment questions. It is unlikely that these questions carried a response bias towards not reporting the objects, as the questions were initially non-directive and gave the participants ample opportunity to describe their experience before explicitly addressing the presence of the objects. There is, of course, a confounding factor in that the questionnaire was administered some minutes after completing the task, which introduced a component of visual memory that was required in reporting the objects but which was not required during the attention task. The subsequent experiments therefore addressed this issue by requiring subjects to complete a secondary task following the attention task. Across many variants of this task, essentially participants were unable to discriminate when the objects were present from when they were absent. In experiment five, they were shown to be unable to discriminate "within-object" trials from "between-object" trials, the critical property subserving the attentional effect. This was important as it may have been the case that participants experienced some perception that the cue and target fell within a single object compared to when they did not, without experiencing a conscious perception of the objects *per se*. The final experiment (six), however, is certainly the most convincing argument that the object were truly invisible to the participants, as in this case the nature of the objects were fully revealed to the participants, and the task of discriminating “objects-present” from “objects-absent” trials is arguably the one which is the least cognitively demanding. The somewhat tricky issue of how to appropriately assess awareness in unseen cueing/priming tasks has therefore been addressed in the present set of experiments by use of a variation of assessment techniques, ranging from subjective reports to the most sensitive confidence-rating signal detection method that is possible.

The question of whether the experiments reported here are truly measuring what is typically described as object-based attention has also been suitably addressed. Importantly, the onset asynchrony between the cue and target (160ms) is typical of the exogenous properties of object-based attention, which require some time-course within the range of 90 and 300ms to be successfully engaged (de-Wit *et al*, 2011). The effect has also been shown to occur (experiment three) when the transient cue is not confounded with a contingency between the object on which it appears and the upcoming target, and the effect is also robust to when the objects appear at unpredictable locations and assume unpredictable orientations (experiments four to six). This automatic and preattentive nature of object-based attention is an important one, and one which has not been suitably studied in previous attempts to dissociate object-based attention from awareness (Chou & Yeh, 2012; Ariga *et al*, 2007).

The presence of the object-based attention effect implies that the 2nd-order contours that separate the figures from their background have been processes by the visual system insofar as they are integrated into a “unique representational entity” (Naber *et al*, 2011, p.6), as the low-level segmentation components of visual perception are not sufficient to warrant such object-related facilitation (de-Wit *et al*, 2009; Naber *et al*, 2011). It is very likely that this object-based attention necessitates the object-construction processes that occur in LOC (de-wit *et al*, 2009a). It is a fair assumption, therefore, that the objects in these experiments are being encoded at a non-conscious level within LOC; an area of the ventral processing stream which is otherwise strongly associated with object awareness (e.g.

Carlson *et al*, 2007). In Carlson *et al*'s study, object-substitution masking was used to render objects invisible; the critical difference between this method and the general one reported here may be that, in the present experiments, the crucial object information (i.e. the location of the contours) is shared between successive frames, albeit with a different phase. Early segmentation processes show signs of phase-invariance (e.g. Norman *et al*, 2011) and it is known that LOC is capable of representing the structure of objects despite gross changes in the information that defines them (Grill-Spector *et al*, 1998). It is possible therefore, that the object-based attention effects observed in the present experiments reflect the processing ability of both low-level segmentation and high-level object-construction in temporally integrating spatially congruent information (i.e. the positions of contours).

7. General discussion

7.1. 2nd-order segmentation

The work in this thesis was motivated by the clear distinction found in the visual system between its respective processing of the edges and regions of a visual scene. This is most notably apparent with the perception of contours and surfaces based on 1st-order stimuli (Paradiso & Nakayama, 1991; Rogers-Ramachandran & Ramachandran, 1998; Breitmeyer, Kafaligönül, Ögmen, Mardon, Todd & Ziegler, 2006); considerably less is known, however, of the perception of those based on 2nd-order stimuli, to which the visual system is also very sensitive. Studies one and two of this thesis specifically set out to develop our understanding of the processes of 2nd-order segmentation (often termed "texture segmentation"), specifically of mean orientation contrast. It is thought that such segmentation is accomplished through a filter-rectify-filter process (see chapter 1 for an introduction, and Bergen, 1991, for a detailed review of such a model). Its three-stage process clearly distinguishes itself from that which underlies 1st-order luminance segmentation, the basis of which can be found in as early as the centre-surround antagonistic structure of the receptive field of a single retinal ganglion cell (Kuffler, 1953; Hartline and Ratliff, 1957; 1958).

Study one explored the extent to which 1st- and 2nd-order segmentation share certain processing characteristics. It is known, for instance, that 1st-order contours can be extracted by the magnocellular subdivision of processing in a rapid, phase-insensitive manner (Saito, Tanaka, Isono, Yasuda & Mikami, 1989). This means there are cells that are sensitive to the presence of a contour irrespective of its sign of contrast. It is known that 2nd-order segmentation is rapid, in contrast to 2nd-order discrimination, but whether or not this temporal processing advantage reflects the selective involvement of phase-insensitive magnocellular processes is uncertain. In two of the experiments in the first study a procedure developed by Rogers-Ramachandran and Ramachandran (1998) was adapted to examine 2nd-order segmentation. Specifically, the orientations of two abutting fields of Gabor patches reversed temporally by 90°. In line with Rogers-Ramachandran and Ramachandran's original findings, subjects could readily pick out the contour (i.e. they could segment) at a much higher frequency (~16Hz) than at which they could discriminate the sign of the orientations (~10Hz). This effect was not found for comparable stimuli that differed in orientation variance (which do not automatically segment), indicating that the effect is specific to the edge-based mechanisms involved in segmentation. The perception of the contour also increased with non-foveal viewing, another hallmark of magnocellular processing, whereas the discrimination of the sign of the contour did not. The finding that 2nd-order segmentation improves with non-foveal viewing is not a new one (Joffe & Scialfa, 1995; Saarinen, Rovamo & Virsu, 1987); the experiment reported here, however, shows that this effect is related specifically to edge-based processes and further develops the dissociation between the two processes and helps our understanding of the respective neural processes that may be involved.

Study two was important because there is little agreement as to the cortical locus of 2nd-order segmentation, despite considerable efforts from studies of cell-recording (e.g. Knierim & van Essen, 1992; Sillito, Grieve, Jones, Cudeiro & Davis, 1995; Nothdurft, Gallant & van Essen, 2000; Lamme, Supèr, & Spekreijse, 1998; Lamme, Rodriguez-Rodriguez & Spekreise, 1999; Hupé, James, Girard & Bullier, 2001; Hegdé & Van Essen, 2000, 2003; Hegdé & Van Essen, 2004; El-Shamayleh and Movshon, 2011), neuroimaging (e.g. Thielscher, Kolle, Neumann, Spitzer & Gron, 2008; Kastner, de Weerd and Ungerleider, 2000; Schira, Fahle, Donner, Kraft & Brandt, 2004; Montaser-Kouhsari, Landy, Heeger & Larsson, 2007; Hallum, Landy and Heeger, 2011), computational modelling (e.g. Thielscher & Neumann, 2007) and lesions in both non-human primates (e.g. Merigan, Nealey & Maunsell, 1993; Merigan, 2000) and humans (e.g. Allen, Humphreys, Colin & Neumann, 2009). Specifically, it was shown in study two that patient MS has a functionally intact 2nd-order segmentation process; his ability to segment targets embedded even in complex "flow-pattern" displays was well above chance, and he performed consistently less accurately when the target's contours were masked (in line with findings from normal observers in study one), implying that he was using the information present at the contours when it was available to him. Perhaps most importantly, however, MS was able to segment when the orientations reversed at a frequency of 12.5Hz, a level at which he was unable to discriminate the regions. This is strong evidence that MS's ability to segment is similar to that of normal observers in that it is rapid and mediated potentially by phase-insensitive channels.

The finding that MS retains a normal capacity to segment based on 2nd-order contrast has some important neuropsychological implications. Some authors, for instance, have theorised that 2nd-order segmentation requires a recurrent flow of activity between mid-level areas (specifically the human homologue of "V4") and low-level areas (V1 and V2), with the final representation occuring in V4 (Thielscher *et al*, 2008; Thielscher & Neumann, 2009), a notion that is supported by evidence from neuroimaging (Kastner *et al*, 2000; Thielscher *et al*, 2008). Lesion studies are invaluable in testing such a theory, and the experiments conducted with MS, whose lesion includes the area labelled as "V4"[12] in related 2nd-order segmentation studies (Thielscher *et al*, 2008; Thielscher & Neumann, 2009; Allen *et al*, 2009), clearly demonstrate that this region is not necessary for 2nd-order segmentation. In which of the remaining intact regions in MS, then, does 2nd-order segmentation take place? A considerable body of evidence indicates that areas V1 (e.g. Lamme *et al,* 1998; 1999; Hupé *et al,* 2001; Hallum *et al,* 2011; Appelbaum, Ales & Norcia, 2012) and V2 (Hegdé & Van Essen, 2000, 2003, 2004; Anzai, Peng & van Essen, 2007; Merigan, Nealey & Maunsell, 1993; Schira *et al*, 2004) exhibit the appropriate sensitivity to 2nd-order contours. The partial damage to area V2 in patient HJA may explain his severe impairment in segmenting a texture-defined bar (Allen *et al*, 2009), and the observation of activity in V4 in fMRI studies may only reflect feedforward projections from this area as well as attention-related effects which are difficult to rule out in imaging experiments; a theory which is supported by at least one imaging study (Schira *et al*, 2004). Similarly, the sometimes observed activity in LOC (Thielscher *et al,* 2008; Larsson *et al*, 2006) is likely only to reflect the area's more general role in form perception and not the processes of segmentation *per se* (Thielscher *et al*, 2008; Merigan, 2000). In conclusion, although we cannot yet be certain whether V1 itself is sufficient or whether V2[13] is necessary, the results from MS quite convincingly show that "V4" is not responsible for segmentation.

---

[12] Chapter 3 includes a note addressing the parity between macaque V4 and human V4

[13] Results from study four of this thesis that show a lower visual field superiority for 2nd-order segmentation may imply a cortical locus at least at the level of V2.

7.2. Perception of orientation variance

As described in the introduction to this thesis, the term "discrimination" when applied to texture perception is often reserved for those instances in which the regions of texture do not segment, typically either through spatial or temporal separation, and so discerning whether two regions are in fact different requires scrutiny. In study one, the difference between segmenting and discriminating regions that differ in their mean orientation was evident, in that observers showed superior performance, in terms of accuracy, reaction times and flicker frequency thresholds, when segmentation was possible. When required to perform a very different type of analysis, however, their performance was quite different. Specifically, when judging differences in orientation variance, they were consistently quicker when the patches of texture were *separated* compared to when they were abutting and showed identical flicker frequency thresholds in the two conditions. There is a clear distinction, therefore, between judgments of mean orientation and those of orientation variance. The distinction is likely to reflect differences in processing requirements; whereas discontinuities in mean orientation, a first moment statistical property, can be detected by local edge-based processes, discontinuities in orientation variance, a second moment statistic, require regional integration. Judgment of mean orientation, therefore, is likely to only involve the summation of activity arising from the orientation-specific cells in primary visual cortex, which may explain MS's generally good performance in the discrimination conditions in the first two experiments of study two. Judgments of orientation variance, however, necessitate an estimation of how much dispersion there is from this particular mean value.

As mentioned, variance discrimination in study one was shown to be made more quickly for separated patches than for abutting patches. Wolfson and Landy (1998) briefly discussed that it might be expected that subjects would perform better on judgments of variance if the patches were separated, suggesting that the separation of the textures serves to clearly delineate the regions to be discerned. The flexible "region of integration" associated with the encoding of spatially distributed orientation information (Dakin & Watt, 1997; Dakin, 2001) may indeed benefit from the clear border definitions offered by the separation of the textures. Dakin's notion of a region of integration implies that there is a dedicated mechanism that is encoding the amount of orientation variance in a stimulus, an idea developed further by Morgan, Chubb and Solomon (2008). Study three set out to determine whether what was being represented by the visual system in the cases of orientation variance discrimination indeed originated from a dedicated processing mechanism that explicitly encoded this second moment regional statistic. The psychophysical technique of adaptation was used to reveal perceptual aftereffects that were specific to high and low levels of orientation variance. Specifically, when subjects adapted to a patch of low orientation variance, a subsequently viewed patch of medium variance was perceived to in fact have a larger degree of variance. The complementary effect following adaptation to high variance was also found. Much like adaptation to other stimulus features (e.g. motion, colour etc.) this reveals that orientation variance is explicitly encoded in the visual system by channels tuned to high and low variance, and that it is the neural activity within each of these channels that determines the perceptual output. This finding adds to others which have shown selective adaptation to a mechanism that explicitly encodes the statistics of a visual image (e.g. density and glossiness). Importantly, however, the results from the same study also imply that the level at which orientation variance is ultimately represented is not dependent on the low-level, local information present in the region. Specifically, the randomisation of the mean orientation throughout the adaptation experiment ensured that adaptation effects to single orientations and mean orientation were ruled out; orientation variance, therefore, is encoded and represented independently of the individual orientations that constitute the texture.

The aftereffects to orientation variance, however, were only specific to a retinotopic reference frame. That is, when observers relocated their gaze between adaptation and test, no aftereffect was observed for stimuli presented in the same spatial location as the adapting stimuli, suggesting that the encoding of orientation variance is not spatiotopic. This is also true of the tilt aftereffect, which follows adaptation to a 1st-order orientation property. Knapen, Rolfs, Wexler & Cavanagh (2010) argue that it would be unlikely to find spatiotopic aftereffects of orientation adaptation, due to the encoding of orientation being carried out in V1; remapping of the adaptation would require that the modified, adapted state of neurons be transmitted horizontally through the brain in any direction depending on the impending eye movement, necessitating a very dense connectivity in lower visual areas for which there seems to be no neurophysiological evidence. It was theorised in study three that aftereffects to variance adaptation may have revealed spatiotopic encoding, due to the large receptive fields that may be required to achieve second moment orientation-based estimations and the spatiotopic properties of visual areas that present such large receptive fields. In study three, however, no evidence for this was found. This may indicate that variance perception is achieved and represented at a much earlier level in the visual system than previously expected. The results from experiment four in study three also develop this theory, as patient MS was shown to be able to discriminate levels of orientation variance. Importantly, MS could not have achieved this on a purely local scale; the randomisation of the mean orientations in the stimuli ensured that an estimate of variance independently of mean orientation for at least two of the three textures was made. In light of recent evidence it may not be surprising that early cortical areas may encode orientation variance; primary visual cortex has been shown to be sensitive to long range contextual effects of orientation (Joo, Boynton & Murray, 2012). Together, the results from these experiments suggest that even a second moment regional statistic of orientation processing may be computed as early as primary visual cortex.

7.3. Cue-invariant segmentation

As information processing in the ventral stream becomes more complex, it also becomes more object-oriented. The neural representation of objects, which is found in lateral occipital complex (LOC; Malach *et al*, 1995; Goodale, Milner, Jakobson & Carey, 1991), for instance, is cue-invariant (Grill-Spector, Kushnir, Edelman, Itzchak & Malach, 1998). This reflects the need to encode objects independently of how they are defined with respect to their environment, thus achieving a degree of perceptual constancy. Study four set out to further our understanding of cue-invariant segmentation, and how extrastriate areas may contribute to its processing. The study first showed that when the feature defining a target region rapidly alternated between one type of 2nd-order feature (orientation) and another (spatial frequency), the segmentation of this target was equivalent to that of standard single-cue segmentation. This may imply neural synergy between the representation of contours by orientation and spatial frequency contrast, and the rapid cue-alternating stimuli offered a novel method with which to explore this aspect of visual processing.

The second aim of study four was to determine whether this cue-invariant aspect of segmentation necessitated additional neural processing beyond the level of striate cortex than single-cue segmentation. This was demonstrated by exploiting the upper-lower visual field asymmetry that is attributed to dedicated cortical processing at the level of V2 and beyond (Horton & Hoyt, 1991; He, Cavanagh & Intrilligator, 1996; Graham, Rico, Offen & Scott, 1999). It was shown that observers were more sensitive to a figure defined by an alternating cue in the lower visual field than in the upper visual field, and that the magnitude of this asymmetry was considerably larger than that of single-cue segmentation. This effect was not present when observers were required to perform the task by discriminating the stimuli (i.e. with black border masks present), indicating that it is an effect specific to the construction of the figure's contours. A likely candidate subserving this process would seem to be LOC; an area which, in addition to its cue-invariance, shows a greater representation of the lower visual field (Sayres & Grill-Spector, 2008). To test this theory, patient DF was tested on her ability to segment on the basis of the single- and alternating-cue stimuli. As DF's lesion includes LOC, it was theorised that she would be selectively impaired in segmenting a figure that was defined by a rapidly alternating cue compared to one defined by a single cue. Her performance, however, was almost identical across the conditions, suggesting that the cortical area responsible for processing the alternating-cue stimulus is not found in LOC. The neural representation of segmentation independently of the type of feature cue, therefore, may be found as early as V2, which has been shown to exhibit some degree of cue invariance (Leventhal, Wang, Schmolesky & Zhou, 1988; Marcar, Raiguel, Xiao & Orban, 2000).

7.4. Segmenting objects as units of attention without awareness

The discussion in the previous section focussed on the importance of cue-invariant representations in higher level visual processing. Cue-invariance represents a midway point between low-level segmentation and the construction of a meaningful perceptual object. Although it was quite clear from the previous set of experiments that LOC is not necessary for at least some forms of cue-invariant processing, it does seem to be necessary for the final construction of object representations; one corollary of which is the ability to select objects as units of attention. Study five set out to determine whether an object-based attention effect could be observed without awareness of the objects, and set out to show this under a variety of experimental procedures, which is highly preferable in measuring the implicit effects of unseen stimuli. The objects were masked from awareness by reversing the sign of an orientation contrast that defined them at a level of 16.7Hz.

In the initial two experiments, the assessment of awareness came only from the participants' report of their subjective experience of the experiment, which was acquired through a series of post-experiment questions. The subsequent experiments developed a more stringent assessment of awareness by requiring subjects to complete a secondary task following the attention task. Across many variants of this task, essentially the results confirmed that participants were extremely unlikely to have had any conscious access to the objects. A number of different procedures were used to measure the subjects' sensitivity, and thus the issue of how to appropriately assess awareness in unseen cueing/priming tasks without overestimating or inaccurately measuring d' (d_a in this case) was addressed, as subjects consistently failed to produce above-chance discrimination scores in the signal detection condition despite numerous different approaches. The results from the experiments, therefore, are unlikely to simply reflect differences in task-demands between the attention and signal detection tasks.

As the benefits conferred by object-based attention necessitate a unique representation of objecthood (Naber, Carlson, Verstaten & Einhäuser, 2011), the object-based attention effect is very likely to necessitate the object-construction processes that occur in LOC (de-wit *et al*, 2009a). It is a fair assumption, therefore, that the objects in these experiments are being encoded at a non-conscious level within LOC; an area of the ventral processing stream which is otherwise strongly associated with object awareness (e.g. Carlson *et al*, 2007). In Carlson *et al*'s study, object-substitution masking was used to render objects invisible and they demonstrated that neural activity within this region correlated with the presence/absence of perception. The critical difference between this method and the general one reported in study five of this thesis may be that, in study five, the crucial object information (i.e. the location of the contours) is shared between successive frames, albeit with a different phase. Early segmentation processes show signs of phase-invariance (e.g. Norman *et al*, 2011) and it is known that LOC is capable of representing the structure of objects despite gross changes in the information that defines them (Grill-Spector *et al*, 1998). It is possible therefore, that the object-based attention effects observed in the present experiments reflect the processing ability of both low-level segmentation and high-level object-construction in temporally integrating spatially congruent information (i.e. the positions of contours). At this point it is not possible to determine whether the associated object representation that is likely to be occurring in LOC is simply operating below a threshold of visual awareness, or whether the activity is instead qualitatively distinct somehow from that of normal conscious representation.

The selection of the objects in this paradigm is clearly driven by the rapid segmentation processes that are capable of signalling a contour despite rapid changes in the sign across the contour. In showing object-based attention without object awareness with such stimuli, these experiments have demonstrated that the transition between this low-level segmentation and high-level object construction is an unconscious one, and that these object representations are even capable of being selectively enhanced by the visual system's attentional processes. This has much broader significance in the field of cognitive psychology, as it is a demonstration of attention selectively enhancing perceptual information that is not limited to a spatial reference frame. Additionally, although the effect of attention on unseen stimuli has been well-studied (e.g. Kentridge, Heywood & Weiskrantz, 1999; Kentridge, Nijboer & Heywood, 2008; Wyart & Tallon-Baudry, 2008), it has been contended on the basis that there has been no parity between the unit of attention and the object of awareness (Mole, 2008). In study five, however, it is clear that there is such parity, and so it strengthens the general claim that attention is not a sufficient precondition for awareness.

The object-based attention effects observed in study four stem from simple segmentation processes; however, the visual system also uses grouping principles to infer the true nature of the environment when complete segmentation information is not immediately available (i.e. when objects may be partially occluded). Consequently, object-based attention effects are still found for partially-occluded objects (Moore, Yantis & Vaughan, 1998). An important step in future research, therefore, will be to determine whether such grouping principles that extend beyond simple segmentation processes impose a limitation on the functionality of object-based attention in the absence of awareness. Furthermore, it also remains to be seen whether a similar dissociation between attention and awareness is found for voluntarily directed object-based attention, although there are clearly more difficulties one can envisage that would be associated with attempting to demonstrate this particular dissociation.

8. References

Adelson, E. (2001). On seeing stuff: The perception of materials by humans and machines. In *Proceedings of the SPIE*, *Vol. 4299*, 1—12.

Adelson, E. H., & Bergen, J. R. (1991). The plenoptic function and the elements of early vision. In M. S. Landy & J. A. Movshon (Eds.), *Computational Models of Visual Processing* (pp. 3-20). Cambridge, MA: MIT Press.

Afraz, S., & Cavanagh, P. (2009). The gender-specific face aftereffect is based in retinotopic not spatiotopic coordinates across several natural image transformations. *Journal of Vision, 9,* 1-17

Allen, H., Humphreys, G., & Bridge, H. (2007). Ventral extra-striate cortical areas are required for optimal orientation averaging. *Vision Research, 47,* 766-775

Allen, H., Humphreys, G., Colin, J., & Newmann, H. (2009). Ventral extra-striate cortical areas are required for human visual texture segmentation. *Journal of Vision, 9,* 1-14

Anzai, A., Peng, X., & Van Essen, D. (2007). Neurons in monkey visual area V2 encode combinations of orientations. *Nature Neuroscience, 10,* 1313-1321

Appelbaum, L., Ales, J., & Norcia, A. (2012). The time course of segmentation and cue-selectivity in the human visual cortex. *PLoS ONE, 7,* e34205

Appelbaum, L., Wade, A., Vildavski, V., Pettet, M., & Norcia, A. (2006). Cue-invariant networks for figure and background processing in human visual cortex. *The Journal of Neuroscience, 26,* 11695-11708

Arcizet, F., Jouffrais, C., & Girard, P. (2008). Natural textures classification in area V4 of the macaque monkey. *Experimental Brain Research, 189,* 109-120

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

Ariga, A., Yokosawa, K., & Ogawa, H. (2007). Object-based attentional selection and awareness of objects. *Visual Cognition, 15,* 685–709

Azzopardi, P., & Cowey, A. (1997). Is blindsight like normal, near-threshold vision? *Proceedings of the National Academy of Science (USA), 94,* 14190-14194

Beason-Held, L., Purpura, K., Krasuski, J., Maisog, J., Daly, E., Mangot, D., *et al.* (1998). Cortical regions involved in visual texture perception: a fMRI study. *Cognitive Brain Research, 7,* 111-118

Beck, J. (1972). Similarity grouping and peripheral discriminability under uncertainty. *American Journal of Psychology, 85,* 1-19

Beck, J. (1982). Textural segmentation. In J. Beck (Ed.), *Organization and representation in perception* (pp. 285-317). Lawrence Erlbaum Associates, Hillsdale, NJ.

Beck, J. (1983). Textural segmentation, 2nd-order statistics, and textural elements. *Biological Cybernetics, 48,* 125-130

Beck, D., & Kastner, S. (2009). Top-down and bottom-up mechanisms in biasing competition in the human brain. *Vision Research, 49,* 1154–1165

Ben-Shahar, O., Scholl, B., & Zucker, S. (2007). Attention, segregation, and textons: Bridging the gap between object-based attention and texton-based segregation. *Vision Research*, *47,* 845-860

Bergen, J. R. (1991). Theories of visual texture perception. In D. Regan (Ed.), *Vision and Visual Dysfunction, vol. 10B*, pp. 114-134. New York: Macmillan.

Bergen, J. R., & Adelson, E. H. (1988). Early vision and texture perception, *Nature*, 333, 363-364.

Bermeitinger, C., Kuhlmann, M., & Wentura, D. (2012). Reading a standing wave: Figure-ground-alternation masking of primes in evaluative priming. *Consciousness & Cognition, 21,* 1109-1121

Biederman, I., & Ju, G. (1988). Surface versus edge –based determinants of visual recognition. *Cognitive Psychology, 20,* 38-64

Blake, A., & Marinos, C. (1990). Shape from texture: estimation, isotropy and moments. *Journal of Artificial Intelligence*, *45,* 323–380

Blakemore, C., Carpenter, R., & Georgeson, M. (1970) Lateral inhibition between orientation detectors in the human visual system. *Nature, 228,* 37-39

Blasdel, G., & Lund, J. (1983). Termination of afferent axons in macaque striate cortex. *Journal of Neuroscience, 3,* 1389-1413

Bovik, A., Clark M., & Geisler, W. (1990). Mulitchannel texture analysis using localized spatial filters. *Pattern Analysis and Machine Intelligence 12,* 55-73

Boycott, B., & Wässle, H. (1991). Morphological classification of bipolar cells of the primate retina. *European Journal of Neuroscience, 3,* 1069-1088

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

Breitmeyer, B., Öğmen, H., Ramon, J., & Chen, J. (2005). Unconscious and conscious priming by forms and their parts. *Visual Cognition, 12,* 720-736

Briand, K. A. (1998). Feature integration and spatial attention: More evidence of a dissociation between endogenous and exogenous orienting. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 1243–1256.

Caelli, T. (1985). Three processing characteristics of visual texture segmentation, *Spatial Vision*, *1*, 19-30.

Caelli, T., & Julesz, B. (1978). On perceptual analyzers underlying visual texture discrimination: Part I, *Biological Cybernetics*, *28,* 167-175.

Caelli, T., Julesz, B., & Gilbert, E. (1978). On perceptual analyzers underlying visual texture discrimination: Part II, *Biological Cybernetics*, *29*, 201-214.

Cant, J., & Goodale, M. (2007). Attention to form or surface properties modulates different regions of human occipitotemporal cortex. *Cerebral Cortex, 17,* 713-731

Cant, J., Arnott, S., & Goodale, M. (2009). fMR-adaptation reveals separate processing regions for the perception of form and texture in the human ventral visual system. *Experimental Brain Research, 192,* 391-405

Caputo, G. (1998). Texture brightness filling-in. *Vision Research, 38, 841-851*

Carey, D. P., Dijkerman, H. C., Murphy, K. J., Goodale, M. A., & Milner, A. D. (2006). Pointing to places and spaces in a patient with visual form agnosia. *Neuropsychologia, 44*, 1584-1594.

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

Carlson, T., Rauschenberger, R., & Verstraten, F., (2007). No representation without awareness in the lateral occipital cortex. *Psychological Science, 18,* 298-302

Carrasco, M. (2011). Visual attention: the past 25 years. *Vision Research, 51,* 1484-1525

Cavina-Pratesi, C., Kentridge, R.W., Heywood, C.A., & Milner, A.D. (2010a). Separate processing of texture and form in the ventral stream: Evidence from fMRI and visual agnosia. *Cerebral Cortex, 20,* 433-446

Cavina-Pratesi, C., Kentridge, R.W., Heywood, C.A., & Milner, A.D. (2010b). Separate channels for processing form, texture, and color: Evidence from fMRI adaptation and visual object agnosia. *Cerebral Cortex, 20,* 2319-2332

Chica, A.B., Lasaponara, S., Chanes, L., Valero-Cabré, A., Doricchi, F., Lupiáñez, J. & Bartolomeo, P. (2011). Spatial attention and conscious perception: the role of endogenous and exogenous orienting. *Attention perception psychophysics*,*73*, 1065-1081

Cho, R. Y., Yang, V., & Hallett, P. E. (2000). Reliability and dimensionality of judgments of visually textured materials. *Perception & Psychophysics, 62,* 735–752.

Chong, S. C., & Treisman, A. (2003). Representation of statistical properties. *Vision Research, 43,* 393–404.

Chou, W., & Yeh, S. (2012). Object-based attention occurs regardless of object awareness. *Psychonomic Bulleting & Review, 19,* 225-231

Chubb, C., Olzak, L., & Derrington, A. (2001). 2nd-order processes in vision: introduction. *Journal of the Optical Society of America A, 18,* 2175-2178

Chubb, C., Sperling, G., & Solomon, J. (1989). Texture interactions determine perceived contrast. *Proceedings of the National Academy of Science (USA). 86,* 9631-9635

Clarke, D., & Sokoloff, L. (1994). Circulation and energy metabolism of the brain. In G. J. Siegel, B. W. Agranoff, R. W. Albers, & P. B. Molinoff (Eds.), *Basic Neurochemistry* (pp. 645–660). New York: Raven.

Corbett, J., Wurnitsch, N., Schwartz, A., & Whitney, D. (2012). An aftereffect of adaptation to mean size. *Visual Cognition, 20,* 211-231

Crespi, S., Biagi, L., d'Avossa, G., Burr, D., Tosetti, M., & Morrone, M. (2011). Spatiotopic of BOLD signal in human visual cortex depends upon spatial attention. *PLoS One, 6,* e21661

Dai, J., & Wang, Y. (2011). Representation of surface luminance and contrast in primary visual cortex. *Cerebral Cortex, 22,* 776-787

Dakin, S. (2001). Information limit on the spatial integration of local orientation signals. *Journal of the Optical Society of America A, 18,* 1016-1026

Dakin, S., & Watt, R. (1997). The computation of orientation statistics from visual texture. *Vision Research, 37,* 3181-3192

Davis, G., Driver, J., Pavani, F., & Shepherd, A. (2000). Reappraising the apparent costs of attending to two separate objects. *Vision Research, 40,* 1323-1332

Dehaene, S., Naccache, L., Cohen, L., Le Bihan, D., Mangin, J., Poline, J., & Rivère, D. (2001). Cerebral mechanisms of word masking and unconscious repetition priming. *Nature Neuroscience, 4,* 752-758

Desimone, R., & Duncan, J. (1995**).** Neural mechanisms of selective visual attention. *Annual Review of Neuroscience, 18,* 193-222

De Valois, R. & De Valois, K. (1988). *Spatial vision*. New York: Oxford University Press

de-Wit, L. H., Cole, G. G., Kentridge, R. W. & Milner, A. D. (2011). The parallel representation of the objects selected by attention*. Journal of Vision, 11,* 1-10

de-Wit, L. H., Kentridge, R. W. & Milner, A. D. (2009a) Object-based attention and visual area LO*. Neuropsychologia, 47*, 1483-1490

de-Wit, L. H. and Kentridge, R. W. and Milner, A. D. (2009b). Shape processing area LO and illusory contours. *Perception., 38,* 1260-1263

Diamond, A. (1953). Foveal simultaneous brightness contrast as a function of inducing- and test-field luminances. *Journal of Experimental Psychology, 45,* 304 – 314

Dow, B. (2002). Orientation and colour columns in monkey visual cortex. *Cerebral Cortex, 12,* 1005-1015

Driver, J., Davis, G., Ricciardelli, P., Kidd, P., Maxwell, E., & Baron-Cohen, S. (1999). Gaze perception triggers reflexive visuospatial orienting. *Visual Cognition, 6,* 509-540.

Duhamel, J., Colby, C., & Goldberg, M. (1992). The updating of the representation of visual space in parietal cortex by intended eye movements. *Science, 255,* 90-92

Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General, 113,* 501-517

Durgin, F. H. (2001). Texture contrast aftereffects are monocular; texture density aftereffects are binocular. *Vision Research, 41,* 2619-2630

Durgin, F. H., & Huk, A. C. (1997). Texture density aftereffects in the perception of artificial and natural textures. *Vision Research, 37,* 3273-3282

Egly, R., Driver, J., & Rafal, R. (1994). Shifting attention between objects and locations: Evidence from normal and parietal lesion subjects. *Journal of Experimental Psychology: General, 123,* 161-177

Ekroll, V., Faul, F., & Niederée, R. (2004). The peculiar nature of simultaneous colour contrast in uniform surrounds. *Vision Research, 44,* 1765-1786

El-Shamayleh, Y., & Movshon, A. (2011). Neuronal response to texture-defined form in macaque visual area V2. *Journal of Neuroscience, 31,* 8543-8555

Elder, J., & Zucker, S. (1998). Evidence for boundary specific grouping. *Vision Research, 38,* 143-152

Eriksen, C., & Yeh, Y. (1985). Allocation of attention in the visual field. *Journal of Experimental Psychology: Human Perception and Performance, 11*, 583–597.

Fahrenfort, J., Scholte, H., & Lamme, V. (2007). Masking disrupts reentrant processing in human visual cortex. *Journal of Cognitive Neuroscience, 19,* 1488-1497

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

Field, D., Hayes, A., & Hess, R. (1993). Contour integration by the human visual system: evidence for a local “association field”. *Vision Research, 33,* 173-193

Findlay, J.M. (2005). Covert attention and saccadic eye movements. In Itti, L., Rees, G. & Tsotsos, J. (Eds.), *Neurobiology of Attention* (pp. 114-117). London New York: Academic Press.

Fleming, R., Dror, R., & Adelson, E. (2003). Real-world illumination and the perception of surface reflectance properties. *Journal of Vision, 3,* 347-368

Fogel, I., & Sagi, D. (1989). Gabor filters as texture discriminator. *Biological Cybernetics, 61,* 103-113

Friesen, C., Ristic, J., & Kingstone, A. (2004). Attentional effects of counterpredictive gaze and arrow cues. *Journal of Experimental Psychology: Human Perception and Performance, 30,* 319-329.

Gardner, J., Merriam, E., Movshon, J., & Heeger, D. (2008). Maps of visual space in human occipital cortex are retinotopic, not spatiotopic. *Journal of Neuroscience, 28,* 3988-3999

Gerrits, H., & Timmerman, G. (1969). The filling-in process in patients with retinal scotomata. *Vision Research, 9,* 439-442

Giordano, A. M., McElree, B., & Carrasco, M. (2009). On the automaticity and flexibility of covert attention: A speed-accuracy trade-off analysis. *Journal of Vision, 9,* 10–31

Golomb, J., & Kanwisher, N. (2011). Higher level visual cortex represents retinotopic, not spatiotopic, object location. *Cerebral Cortex,* doi: 10.1093/cercor/bhr357

Golomb, J., Nguyen-Phuc, A., Mazer, A., McCarthy, G., & Chun, M. (2010). Attentional facilitation throughout human cortex lingers in retinotopic coordinates after eye movements. *Journal of Neuroscience, 30,* 10493-10506

Goodale, M., Milner, A. D., Jakobson, L., Carey, D. P. (1991). A neurological dissociation between perceiving objects and grasping them. *Nature, 349*, 154–156.

Graham, N. (1989). *Visual Pattern Analyzers,* New York: Oxford University Press

Graham, N. (1991). Complex channels, early local nonlinearities, and normalization in perceived texture segregation. In M. S. Landy & J. A. Movshon (Eds.) *Computational Models of Visual Processing* (pp. 291-301). Cambridge, MA: MIT Press.

Graham, N. (1992). Breaking the visual stimulus into parts. *Current Directions in Psychological Science*, *1*, 55-61.

Graham, N., Rico, M., Offen, S., & Scott, W. (1999). Texture segregation shows only a very small lower-hemifield advantage. *Vision Research*, *39,* 1171-1175

Grill-Spector, K., Kushnir, T., Edelman, S., Itzchak, Y., & Malach, R. (1998). Cue invariant activation in object-related areas of the human occipital lobe. *Neuron, 21,* 191-202

Grill-Spector, K., & Malach, R. (2001). fMR-adaptation; a tool for studying the functional properties of human cortical neurons. *Acta Psychologia (Amst), 107*, 293-321

Hallum, L., Landy, M., & Heeger, D. (2011). Human primary visual cortex (V1) is selective for 2nd-order spatial frequency. *Journal of Neurophysiology, 105,* 2121-2131

Hanazawa, A., & Komatsu, H. (2001). Influence of the direction of elemental luminance gradients on the responses of V4 cells to textured surfaces. *Journal of Neuroscience, 21,* 4490-4497

Hartline, F., & Ratliff, F. (1957). Inhibitory interaction of receptor units in the eye of the limulus. *Journal of General Physiology, 40,* 357-376

Hartline, F., & Ratliff, F. (1958). Spatial summation of inhibitory influences in the eye of the limulus, and the mutual interaction of receptor units. *Journal of General Physiology, 41,* 1049-1066

Harvey, L. (2002). RScorePlus, v 5.6.1. Parameter estimation of signal detection models. Available at http://psych.colorado.edu/~lharvey/

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

Hegdé, J., & Van Essen, D. (2000). Selectivity for complex shapes in primate visual area V2. *Journal of Neuroscience, 20,* RC61-66

Hegdé, J., & Van Essen, D. (2003). Strategies of shape representation in macaque visual area V2. Selectivity for complex shapes in primate visual area V2. *Visual Neuroscience, 20*, 313-328

Hegdé, J., & Van Essen, D. (2004). Temporal dynamics of shape representation in macaque visual area V2. *Journal of Neurophysiology, 92*, 3030-3042

Heinemann, E. (1955). Simultaneous brightness induction as a function of inducing and test-field luminances. *Journal of Experimental Psychology, 50***,** 89–96

Heinrich, S., Andrés, M., & Bach, M. (2007). Attention and visual texture segregation. *Journal of Vision, 7,* 1-10

Hendry, S., & Reid, R. C. (2000). The koniocellular pathway in primate vision. *Annual Review of Neuroscience, 23,* 127-153

Hendry, S., & Yoshioka, T. (1994). A neurochemically distinct third channel in the macaque dorsal lateral geniculate nucleus. *Science, 264,* 575-577

Heywood, C. A., & Cowey, A. (1987). On the role of cortical area V4 in the discrimination of hue and pattern in macaque monkeys. *Journal of Neuroscience, 7,* 2601-2617

Heywood, C., Cowey, A., & Newcombe, F. (1991). Chromatic discrimination in a cortically colour blind observer. *European Journal of Neuroscience, 3,* 802-812

Heywood, C., Cowey, A., & Newcombe, F. (1994). The role of the parvocellular (P) and magnocellular (M) pathways in cerebral chromatiopsia. *Brain, 117,* 245-254

Heywood, C. A., Gadotti, A., & Cowey, A. (1992). Cortical area V4 and its role in the perception of colour. *Journal of Neuroscience, 12,* 4056-4065

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

Hofman, M., & Hallett, P. (1993). Texture segregation based on two-dimensional relative phase differences in composite sine-wave grating patterns. *Vision Research, 33,* 221-234

Holcomb, A. (2009). Seeing slow and seeing fast: two limits on perception. *Trends in Cognitive Science, 13,* 216-221

Horton, J., & Hoyt, W. (1991). Quadrantic visual field defects*.* *Brain, 114,* 1703-1718

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

Hubel, D., & Wiesel, T. (1959). Receptive fields of single neurones in the cat's striate cortex. *Journal of Physiology, 148,* 574-591

Hubel, D., & Wiesel, T. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology, 160,* 106-154

Hubel, D., & Wiesel, T. (1972). Laminar and columnar distribution of geniculo-cortical fibers in the macaque monkey. *Journal of Comparative Neurology, 146,* 421-450

Humphreys, G., Cinel, C., Wolfe, J., Olson, A., & Klempen, N. (2000). Fractitioning the binding process: neuropsychological evidence distinguishing binding of form from binding of surface features. *Vision Research, 40,* 1569-1596

Hupé, J., James, A., Girard, P. & Bullier, J. (2001). Response modulation by static texture surround in area V1 of the macaque monkey do not depend on feedback connections from V2. *Journal of Neurophysiology, 85,* 146-163

Hurlbert, A. & Wolf, K. (2004). Color contrast: a contributory mechanism to color constancy. In C. A. Heywood, A. D. Milner & C. Blakemore (Eds.), *The roots of visual awareness: a festschrift in honour of Alan Cowey* (pp. 147-160). London: Elsevier

James, W. (1890). *The Principles of Psychology*. London: Macmillan

James, T. W., Culham, J., Humphrey, G. K., Milner, A. D., & Goodale, M. A. (2003). Ventral occipital lesions impair object recognition but not object-directed grasping: an fMRI study. *Brain, 126*, 2463-2475

Jans, B., Peters, J., & De Weerd, P. (2010). Visual spatial attention to multiple locations at once: The jury is still out. *Psychological Review, 117,* 637–684

Joffe, K., & Scialfa, C. (1995). Texture segmentation as a function of eccentricity, spatial frequency and target size. *Spatial Vision, 9,* 325-342

Johnson, A., & Baker, C. (2004). First- and second-order information in natural images: a filter-based approach to image statistics. *Journal of the Optical Society of America: Optics, Image Science and Vision, 21,* 913-925

Joo, S. J., Boynton, G., & Murray, S. (2012). Long-range pattern-dependent contextual effects in early human visual cortex. *Current Biology, 9,* 781-786

Juan, C., Shorter-Jacobi, S., & Schall, J. (2004). Dissociation of spatial attention and saccadic preparation. *Proceedings of the Royal Society B, 101,* 15541-15544

Julesz, B. (1962). Visual Pattern Discrimination. *IRE Transactions on Information Theory, 8,* 84–92

Julesz, B. (1981). Textons, the elements of texture perception, and their interactions. *Nature*, *290*, 91-97.

Julesz, B., Gilbert, E., Shepp, L., & Frisch, H. (1973). Inability of humans to discriminate between visual textures that agree in 2nd-order statistics – revisited. *Perception*, *2,* 391-405.

Julesz, B., Gilbert, E., & Victor, J. (1978). Visual discrimination of textures with identical third-order statistics. *Biological Cybernetics*, *31*, 137-140.

Kanai, R., Tsuchiya, N., & Verstraten, F. (2006). The scope and limits of top–down attention in unconscious visual processing. Current Biology, *16, 2332–2336*.

Kass, M., & Witkin, A. (1987). Analyzing oriented patterns. *Computer Vision, Graphics and Image Processing, 37,* 362-385

Kastner, S., De Weerd, P., & Ungerleider, L. (2000). Texture segregation in the human visual cortex: a functional MRI study. *Journal of Neurophysiology, 83,* 2453-2457

Kastner, S., Nothdurft, H., & Pigarev, I. (1997). Neuronal correlates of pop-out in cat striate cortex. *Vision Research, 4,* 371-376

Kastner, S., Nothdurft, H., & Pigarev, I. (1999). Neuronal responses to orientation and motion contrast in cat striate cortex. *Visual Neuroscience, 16,* 587-600

Kentridge R.W. (2011). Attention without awareness: A brief review. In C. Mole, D. Smithies & W. Wu (Eds.), *Attention: Philosophical and Psychological Essays* (pp. 228-246). Oxford: Oxford University Press.

Kentridge, R. W., Heywood, C. A., & Cowey, A. (2004). Chromatic edges, surfaces and constancies in cerebral achromatopsia. *Neuropsychologia, 42,* 821-830

Kentridge, R.W., Heywood, C.A., & Weiskrantz, L. (1999). Attention without awareness in blindsight. *Proceedings of the Royal Society B, 266,* 1805-1811

Kentridge, R.W., Heywood, C.A. & Weiskrantz, L. (2004) Spatial attention speeds discrimination without awareness in blindsight. *Neuropsychologia*, *42*, 831-835.

Kentridge, R. W., Heywood, C. A., & Weiskrantz, L. (2007). Color contrast processing in human striate cortex. *Proceedings of the National Academy of Sciences (USA), 104,* 15129-15131

Kentridge, R.W., Nijboer, T.C.W. & Heywood, C.A. (2008) Attended but unseen: Visual attention is not sufficient for visual awareness. *Neuropsychologia 46, 864-869*

Kim, J., & Anderson, B. (2010). Image statistics and the perception of surface gloss and lightness. *Journal of Vision, 10,* 1-17

Kim, J., & Zatorre, R. (2011). Tactile-auditory shape learning engages the lateral occipital complex. *Journal of Neuroscience, 31,* 7848-7856

Kimchi, R., & Peterson, M. (2008). Figure-ground segmentation can occur without attention. Psychological *Science, 19,* 660-668

Kingdom, F. (1997). Simultaneous contrast: the legacies of Hering and Helmholtz. *Perception*, *26*, 673-677

Knapen, T., Rolfs, M., Wexler, M., & Cavanagh, P. (2009). The reference frame of the tilt aftereffect. *Journal of Vision, 10,* 1-13

Knierim, J., & Van Essen, D. (1992). Neuronal responses to static texture patterns in area V1 of the alert macaque monkey. *Journal of Neurophysiology, 67,* 961-980

Knutsson, H., & Granlund, G. (1983). Texture analysis using two-dimensional quadrature Filters. *Proceedings of the IEEE Computer Society Workshop on Computer Architecture forPattern Analysis and Image Database Management*, Silver Spring, MD: IEEE Computer Society, pp. 206-213

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

Koenderink, J., & van Doorn, A. (1996). Illuminance texture due to surface mesostructure. *Journal of the Optical Society of America A, 13,* 452-463

Kolb, F., & Braun, J. (1995). Blindisght in normal observers. *Nature, 377,* 336-338

Kolb, H., & Dekorver, L. (1991). Midget ganglion cells of the parafovea of the human retina: a study by selective microscopy and serial section reconstructions. *Journal of Comparative Neurology, 303,* 617-636

Kolb, H., & Marshak, D. (2003). The midget pathways of the primate retina. *Documenta Opthalmologica, 106,* 67-81

Kolb, H., Linberg, K., & Fisher, S. (1992). Neurons in the human retina: a Golgi study. *Journal of Comparative Neurology, 318,* 147-187

Kourtzi, Z., & Kanwisher, N. (2001). Representation of perceived object shape by the human lateral occipital complex. *Science, 293,* 1506-1509

Kowler, E. (2011). Eye movements: The past 25 years. *Vision Research, 51,* 1457–1483

Krauskopf, J. (1963). Effect of retinal image stabilization on the appearance of heterochromatic targets. *Journal of the Optical Society of America, 53,* 741-741

Krekelberg, B., Boynton, G., & Van Wezel, R. (2006). Adaptation: from single cells to BOLD signals. *Trends in Neuroscience, 29,* 250-256

Kuffler, S. (1953). Discharge patterns and functional organization of mammalian retina. *Journal of Neurophysiology, 16,* 37-68

Lambert, A., Naikar, N., McLachlan, K., & Aitken, V. (1999). A new component of visual orienting: Implicit effects of peripheral information and subthreshold cues on covert attention. *Journal of Experimental Psychology: Human Perception and Performance, 25,* 321–340.

Lambert, A., & Sumich, A. (1996). Spatial orienting controlled without awareness: A semantically based implicit learning effect. *Quarterly Journal of Experimental Psychology, 49A,* 490–518.

Lamme, V. (1995). The neurophysiology of figure-ground segregation in primary visual cortex. *Journal of Neuroscience, 15,* 1605-1615

Lamme, V. (2003). Why visual attention and awareness are different. *Trends in Cognitive Neuroscience, 7,* 12-18

Lamme, V., 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,* 406-413

Lamme, V., Supèr, H., & Spekreijse, H. (1998). Feedforward, horizontal, and feedback processing in the visual cortex. *Current Opinion in Neurobiology, 8,* 529-535

Landy, M. S., & Bergen, J. R. (1991). Texture segregation and orientation gradient. *Vision Research 31*, 679-691

Landy, M. S., & Graham, N. (2004). Visual perception of texture. In L. Chalupa & J. Werner (Eds.), The visual neurosciences (pp. 1106–1118). Cambridge, MA: MIT Press.

Landy, M. S., & Kojima, H. (2001). Ideal cue combination for localizing texture-defined edges. *Journal of the Optical Society of America A, 18,* 2307-2320

Larimer, J., & Piantanida, T. (1988). The impact of boundaries on color: Stabilized image studies. *Society of Photo-Optical Instrumentation Engineers, 901,* 241-247

Larsson, J., Landy, M. S., & Heeger, D. (2005). Orientation-selective adaptation to first- and 2nd-order patterns in human visual cortex. *Journal of Neurophysiology, 95,* 862- 881

Lee, H., Kramer, A., Mozer, M., & Vecera, S. (2012). Object-based control of attention is sensitive to recent experience. *Journal of Experimental Psychology: Human Perception and Performance, 38,* 314-325

Leventhal, A., Wang, Y., Schmolesky, M., & Zhou, Y. (1998). Neural correlates of boundary perception. *Visual Neuroscience, 15,* 1107-1118

Levine, M., & McAnany, J. (2005). The relative capabilities of the upper and lower visual hemifields. *Vision Research, 45,* 2820-2830

Livingstone, M., & Hubel, D. (1982). Thalamic inputs to cytochrome oxidase-rich regions in monkey visual cortex. *Proceedings of the National Academy of Sciences (USA), 79,* 6098-6101

Livingstone, M., & Hubel, D. (1988). Segregation of form, color, movement, and depth: anatomy, physiology, and perception. *Science, 240,* 740-479

Liu, T., Heeger, D., & Carrasco, M. (2006). Neural correlates of the visual vertical meridian asymmetry. *Journal of Vision, 6,* 1294-1306

Liu, T., Stevens, S. T., & Carrasco, M. (2007). Comparing the time course and efficacy of spatial and feature-based attention. *Vision Research, 47,* 108–113.

Lu, S., Cai, Y., Shen, M., Zhou, Y., & Han, S. (2012). Alerting and orienting of attention without visual awareness. *Consciousness and Cognition, 21,* 928-938

Lu, Z. L., & Dosher, B. A. (1998). External noise distinguishes attention mechanisms. *Vision Research, 38*, 1183–1198.

Lu, Z. L., & Dosher, B. A. (2000). Spatial attention: Different mechanisms for central and peripheral temporal precues? *Journal of Experimental Psychology: Human Perception and Performance, 26,* 1534–1548

Lu, H., & Roe, A. (2008). Functional organizations of color domains in V1 and V2 of macaque monkey revealed by optical imaging. *Cerebral Cortex, 18,* 516-533

MacEvoy, S., & Epstein, R. (2007). Position selectivity in scene- and object-responsive occipitotemporal regions. *Journal of Neurophysiology, 98,* 2089-2098

Mack, A., & Rock, I. (1998). Inattentional Blindness. Cambridge, MA: MIT Press

Macknik, S., & Martinez-Conde, S. (2004). The spatial and temporal effects of lateral inhibitory networks and their relevance to the visibility of spatiotemporal edges. *Neurocomputing, 58-60,* 775-782

Malach, R., Reppas, J., Kwong, K.K., Jiang, H., Kennedy, W.A., Ledden, P.J., *et al.* (1995). Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proceedings of the National Academy of Sciences of the United States of America, 92,* 8135–8138.

Malik, J., & Perona, P. (1990). Preattentive texture discrimination with early vision mechanisms, *Journal of the Optical Society of America A*, *7*, 923-932

Mante, V., Bronin, V., Frazor, R., Geisler, W., & Carandini M. (2005). Independence of gain control mechanisms in early visual system matches the statistics of natural images. *Nature Neuroscience, 8,* 1690–97

Marcar, V., Raiguel, S., Xiao, D., & Orban, G. (2000). Processing of kinetically defined boundaries in areas V1 and V2 of the macaque monkey. *Journal of Neurophysiology, 84,* 2786-2798.

Mareschal, I., & Baker, C. (1998). A cortical locus for the processing of contrast-defined contours. *Nature Neuroscience, 1,* 150-154

Martinez, A., Teder-Salejarvi, W., Vazquez, M., Molholm, S., Foxe, J. J., Javitt, D. C., et al. (2006). Objects are highlighted by spatial attention. *Journal of Cognitive Neuroscience, 18*(2), 298-310.

McCormick, P.A. (1997) Orienting attention without awareness. *Journal of Experimental Psychology: Human Perception and Performance, 23,* 168–180.

McKee, D., Christie, J., & Klein, R. (2007). On the uniqueness of attentional capture by uninformative gaze cues: Facilitation interacts with the Simon effect and is rarely followed by IOR. *Canadian Journal of Experimental Psychology, 61,* 293-303.

McKeefry, D., & Zeki, S. (1997). The position and topography of the human color centre as revealed by functional magnetic resognance imaging. *Brain, 120,* 2229-2242

Melcher, D. (2005). Spatiotopic transfer of visual-form adaptation across saccadic eye movements. *Current Biology, 15,* 1745-1748

Melcher, D., & Colby, C. (2008). Trans-saccadic perception. *Trends in Cognitive Science, 12,* 466-473

Merigan, W. (2000). Cortical area V4 is critical for certain texture discriminations, but this effect is not dependent on attention. *Visual Neuroscience, 17,* 949-958

Merigan,W., Nealey, T., & Maunsell, J. (1993). Visual effects of lesions of cortical area V2 in macaques. *Journal of Neuroscience, 13*, 3180–3191.

Merriam, E., Genovese, C., & Colby, C. (2007). Remapping in human visual cortex. *Journal of Neurophysiology, 97,* 1738-1755

Milner, A. D., & Goodale, M. A. (1995) *The Visual Brain in Action*. Oxford: Oxford University Press.

Milner, A. D., Perrett, D. I., Johnston, R. S., Benson, P. J., Jordan, T. R., Heeley, D. W., et al. (1991). Perception and action in visual form agnosia. *Brain, 114*, 405-428.

Mole, C. (2008). Attention and Consciousness. *Journal of Consciousness Studies, 15, 86-104*

de Monasterio, F., & Gouras, P. (1975). Functional properties of ganglion cells of the rhesus monkey retina. *Journal of Physiology, 251,* 167-195

Montaser-Kouhsari, L., Landy, M. S., Heeger, D., & Larsson, J. (2007). Orientation-selective adaptation to illusory contours in human visual cortex. *Journal of Neuroscience, 28,* 2186-2195

Moore, C., Yantis, S., & Vaughan, B. (1998). Object-based visual selection: Evidence from perceptual completion. *Psychological Science, 9,* 104-110

Moran, J., & Desimone, R. (1985). Selective attention gates visual processing in the extrastriate cortex. *Science, 229,* 782–784

Morgan, M., Chubb, C., & Solomon, J. (2008). A 'dipper' function for texture discrimination based on orientation variance. *Journal of Vision, 8,* 1-8

Morgan, M., Mareschal, I., Chubb, C., & Solomon, J. (2012). Perceived pattern irregularity computed as a summary statistic: Implications for camouflage. *Proceedings of the Royal Society of London. Series B, Biological Sciences, 279,* 2754-2760

Motoyoshi, I. (1999). Texture filling-in and texture segregation revealed by transient masking. *Vision Research, 7,* 1285-1291

Motoyoshi, I., & Nishida, S. (2001). Visual response saturation in response to orientation contrast in the presence of texture boundary. *Journal of the Optical Society of America A, 18,* 2209-2219.

Naber, M., Carlson, T., Verstraten, F., & Einhäuser, W. (2011). Perceptual benefits of objecthood. *Journal of Vision, 11,* 1-9

Nachmias, J., & Sansbury, R. (1974). Grating contrast: Discrimination may be better than detection. *Vision Research, 14,* 1039-1042.

Nakamura, K., & Colby, C. (2002). Updating the visual representation in monkey striate and extrastriate cortex during saccades. *Proceedings of the National Academy of Science (USA), 99,* 4026-4031

Nakayama, K., & Martini, P. (2011). Situating visual search. *Vision Research, 51,* 1526–1537

Nishida, S., & Shinya, M. (1998). Use of image-based information in judgments of surface-reflectance properties. *Journal of the Optical Society of America A, 15,* 2951-2965

Norman, L. J., Heywood, C. A., & Kentridge, R. W. (2011). Contrasting the processes of texture segmentation and discrimination using static and phase-reversing stimuli. *Vision Research, 51,* 2039-2047

Nothdurft, H. (1991). Texture segmentation and pop-out from orientation contrast. *Vision Research, 31*, 1073–1078.

Nothdurft, H., Gallant, J., Van Essen, D. (1999). Response modulation by texture surround in primate area V1: Correlates of 'popout' under anesthesia. *Visual Neuroscience, 16,* 15-34

Nothdurft, H., Gallant, J., Van Essen, D. (2000). Response profiles to texture border patterns in area V1. *Visual Neuroscience, 17,* 421-436.

Palmer, S., & Rock, I. (1994). Rethinking perceptual organization: the role of unform connectedness. *Psychonomic Bulletin & Review, 1,* 29-55

Paradiso, M., & Nakayama, K. (1991). Brightness perception and filling-in. *Vision Research, 31,* 1221-1236

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

Pasley, B., Mayes, L., & Schultz, R. (2004). Subcortical discrimination of unperceived objects during binocular rivalry. *Neuron, 42,* 163-172

Perry, V., Oehler, R., & Cowey, A. (1984). Retinal ganglion cells that project to the dorsal lateral geniculate nucleus in the macaque monkey. *Neuroscience, 12,* 1101-1123

Peuskens, H., Claeys, K., Todd, J., Norman, J., Van Hecke, P., Orban, G. (2004). Attention to 3-D shape, 3-D motion, and texture in 3-D structure from motion displays. *Journal of Cognitive Neuroscience, 16*, 665--682.

Pitts, M., Martinez, A., & Hillyard, S. (2012). Visual processing of contour patterns under conditions of inattentional blindness. *Journal of Cognitive Neuroscience, 24,* 287-303

Pont, S., & te Pas, S. (2006). Material-illumination ambiguities and the perception of solid objects. *Perception, 35,* 1331-1350

Portilla, J., & Simoncelli, E. (2000). A parametric texture model based on joint statistics of complex wavelet coefficients, *International Journal of Computer Vision*, *40,* 49-71

Posner, M. (1980). Orienting of Attention. *Quarterly Journal of Experimental Psychology, 32,* 3-25

Posner, M. I., & Cohen, Y. (1984). Components of visual orienting. In H. Bouma & D.Bonwhuis (Eds.), *Attention and performance X: Control of language processes* (pp. 551–556). Hillsdale, NJ: Erlbaum

Puce, A., Allison, T., Asgari, M., Gore, G., & McCarthy, G. (1996). Differential sensitivity of human visual cortex to faces, letterstrings, and textures: a functional magnetic resonance imaging study. *Journal of Neuroscience, 16,* 5205-5215

Ramachandran, V., & Gregory, R. (1991). Perceptual filling in of artificially induced scotomoas in human vision. *Nature, 350,* 699-702

Ramsden, B., Hung, C., & Roe, A. (2001). Real and illusory contour processing in area V1 of the primate: a cortical balancing act. *Cerebral Cortex, 11,* 648-665

Ratliff, F. (1965). Mach bands: quantitative studies on neural networks in the retina. Holden-Day, San Francisco.

Reingold, E., & Merikle, P. (1988). Using direct and indirect measures to study perception without awareness. *Attention, Perception & Psychophysics, 44,* 563-575

Remington, R. W., Johnston, J. C., & Yantis, S. (1992). Involuntary attentional capture by abrupt onsets. *Perception & Psychophysics, 51,* 279–290.

Reppa, I., Schmidt, W.C. & Leek, E.C. (in press). Successes and failures in producing attentional object-based cueing effects. *Attention, Perception & Psychophysics*, 74 (1), 43-69

Reynolds, J. H., Chelazzi, L., & Desimone, R. (1999). Competitive mechanisms subserve attention in macaque areas V2 and V4. *Journal of Neuroscience, 19,* 1736–1753

Riopelle, A., & Bevan, W. (1952). The distribution of scotopic sensitivity in human vision. *American Journal of Psychology, 66,* 73-80

Ristic, J., & Kingstone, A. (2006). Attention to arrows: Pointing to a new direction. *Quarterly Journal of Experimental Psychology, 59,* 1921-1930

Rivest, J., & Cavanagh, P. (1996). Localizing contours defined by more than one attribute. *Vision Research, 36,* 53-66

Rogers-Ramachandran, D., & Ramachandran, V. (1998). Psychophysical evidence for boundary and surface systems in human vision. *Vision Research, 38,* 17-77

Romani, A., Caputo, G., Callieco, R., Schintone, E., & Cosi, V. (1999). Edge detection and surface 'filling in' as shown by texture visual evoked potentials. *Clinical Neurophysiology, 110,* 86-91

Rubin, N., Nakayama, K., & Shapley, R. (1996). Enhanced perception of illusory contours in the lower versus upper visual field. *Science, 271,* 651-653

Saarela, T., & Landy, M. S. (2012). Combination of texture and color cues in visual segmentation. *Vision Research, 58,* 59-67

Saarinen, J., Rovamo, J., & Virsu, V. (1987). Texture discrimination at different eccentricities. *Journal of the Optical Society of America A, 4,* 1699-1703

Sagi, D., & Hochstein, S. (1985). Lateral inhibition between spatially adjacent spatial-frequency channels? *Attention, Perception & Psychophysics, 37,* 315-322

Saito, H., Tanaka, K., Isono, H., Yasuda, M., & Mikami, A. (1989). Directionally selective response of cells in the middle temporal area (MT) of the macaque monkey to the movement of equiluminous opponent color stimuli. *Experimental Brain Research, 75,* 1-14

Sayres, R., & Grill-Spector, K. (2008). Relating retinotopic and object-selective responses in human lateral occipital cortex. *Journal of Neurophysiology, 100,* 249-267

Schiller, P., & Logothetis, N. (1990). The color-opponent and broad-band channels of the primate visual system. *Trends in Neurosciences, 13,* 392-398

Schiller, P., & Malpeli, J. (1978). Functional specificity of lateral geniculate nucleus laminae of the rhesus monkey. *Journal of Neurophysiology, 41,* 788-797

Schira, M., Fahle, M., Donner, T., Kraft, A., & Brandt, S. (2004). Differential contribution of early visual areas to the perceptual process of contour processing. *Journal of Neurophysiology, 91,* 1716-1721

Schofield, A. (2000). What does second-order vision see in an image? *Perception, 29,* 1071-1086

Self, M., & Zeki, S. (2005). The integration of colour and motion by the human visual brain. *Cerebral Cortex, 15,* 1270-1279

Sereno, M., Dale, A., Reppas., J., Kwong, K., Belliveau, J., Brady, T., *et al.* (1995). Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science, 268,* 889-893

Shapley, R., & Perry, H. (1986). Cat and monkey retinal ganglion cells and their visual functional roles. *Trends in Neurosciences, 9,* 229-235

Sheinberg, D., & Logothetis, N. (1997). The role of temporal cortical areas in perceptual organization. *Proceedings of the National Academy of Sciences (USA), 94,* 3408-3413

Shpaner, M., Murray, M., & Foxe, J. (2009). Early processing in the human lateral occipital complex is highly responsive to illusory contours but not to salient regions. *European Journal of Neuroscience, 30,* 20128-2028

Sillito, A., Grieve, K., Jones, H., Cudeiro, J., & Davis, J. (1995). Visual cortical mechanisms detecting focal orientation discontinuities. *Nature, 378,* 492-496

Skottun, B., & Skoyles, J. (2006). The use of phantom contours to isolate magnocellular and parvocellular responses. *International Journal of Neuroscience, 116,* 315-320

Smith, S., Wenderoth, P., & van der Zwan, R. (2001). Orientation processing mechanisms revealed by the plaid tilt illusion. *Vision Research, 41,* 483-494

Solomon, J. (2010). Visual discrimination of orientation statistics in crowded and uncrowded arrays. *Journal of Vision, 10,* 1-16

Sperling, A., Lu, Z., Manis, F., & Seidenberg, M. (2003). Selective magnocellular deficits in dyslexia: a “phantom contour” study. *Neuropsychologia, 41,* 1422-4129

Sterzer, P., Haynes, J., & Rees, G. (2008). Fine-scale activity patterns in high-level visual areas encode the category of invisible objects. *Journal of Vision, 8, 1-12*

Sterzer, P., Jalkanen, L., & Rees, G. (2009). Electromagnetic responses to invisible face stimuli during binocular suppression. *NeuroImage, 46,* 803-808

Su, Y., He, Z., & Ooi, T. (2011). Seeing grating-textured surfaces begins at the border. *Journal of Vision, 11,* 1-14

Sumner, P., Tsai, P.-C., Yu, K., & Nachev, P. (2006). Attentional modulation of sensorimotor processes in the absence of perceptual awareness. *Proceedings of the National Academy of Sciences (USA)*, *103*, 10520–10525.

Sutter, A., Beck, J., & Graham, N. (1989). Contrast and spatial variables in texture segregation: Testing a simple spatial-frequency channels model. *Perception & Psychophysics, 46,* 312-332

Thielscher, A., Kolle, M., Neumann, H., Spitzer, M., & Gron, G. (2008). Texture segmentation in human perception: a combined modelling and fMRI study. *Neuroscience, 151,* 730-736

Thielscher, A., & Neumann, H. (2003). Neural mechanisms of cortico-cortical interaction in texture boundary detection: a modeling approach. *Neuroscience, 122,* 921–939.

Thielscher, A., & Neumann, H. (2005). Neural mechanisms of human texture processing: Texture boundary detection and visual search. *Spatial Vision, 18,* 227–257.

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.

Troncoso, X., Macknik, S., & Martinez-Conde, S. (2011). The first steps of vision: anatomy, physiology and perception in the normal retina, lateral geniculate nucleus and early visual cortical areas. In G. Dagnelie (Ed.), *Visual Prosthetics: Anatomy, Physiology, Bioengineering, Psychophysics, Rehabilitation*, New York, Springer.

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

Tucker, T., & Fitzpatrick, D. (2006). Luminance-evoked inhibition in primary visual cortex: a transient veto of simultaneous and ongoing response. *Journal of Neuroscience, 26,* 13537-13547

Turi, M., & Burr, D. (2012). Spatiotopic perceptual maps in humans: evidence from motion adaptation. *Proceedings of the Royal Society B,* doi: 10.1098/rspb.2012.0637

Turner, M. (1986). Texture discrimination by Gabor functions. *Biological Cybernetics, 55,* 71-82

Tyler, C., Likova, L., & Wade, A. (2004). Properties of object processing in lateral occipital cortex [Abstract]. *Journal of Vision, 4, 91*

Vermeiren, A., & Cleeremens, A. (2012). The validity of d' measures. *PLoS ONE, 7(2): e31595*

Verweij, J., Dacey, D., Peterson, B., & Buck, S. (1999). Sensitivity and dynamics of rod signals in H1 horizontal cells of the macaque monkey retina. *Vision Research, 39,* 3662-3672

Victor, J., & Conte, M. (2004). Visual working memory for image statistics. *Vision Research, 44,* 541-556

von der Heydt, R., & Peterhans, E. (1989) Mechanisms of contour perception in monkey visual cortex. I. Lines of pattern discontinuity. *Journal of Neuroscience, 9,* 1731-1748.

Vorberg, D., Mattler, U., Heinecke., A., Schmidt, T., & Schwarzbach, J. (2003). Different time courses for visual perception and action priming. *Proceedings of the National Academy of Science (USA), 100,* 6275-6280

Walsh, V., Butler, S., Carden, D., & Kulikowski, J. (1993). The effects of V4 lesions on the visual behaviour of macaques: hue discrimination and colour-constancy. *Behavioural Brain Research, 53,* 51-62

Watanabe, M., & Rodieck, R. (1989). Parasol and midget ganglion cells of the primate retina. *Journal of Comparative Neurology, 289,* 434-454

Watt, R. J., & Morgan, M. J. (1983). The recognition and representation of edge blur: evidence for spatial primitives in human vision. *Vision Research, 23,* 1457-1177.

Weiskrantz, L. (1986). *Blindsight.* Oxford: Oxford University Press.

Wenderoth, P., & Wiese, M. (2008). Retinotopic encoding of the direction aftereffect. *Vision Research, 48,* 1949-1954

Westheimer, G. (1990). Simultaneous orientation contrast for lines in the human fovea. *Vision Research, 30,* 1913-1921

Williams, D. W., & Sekuler, R. (1984). Coherent global motion percepts from stochastic local motions. *Vision Research, 24,* 55–62

Witkin, A. (1981). Recovering surface shape from texture. *Artificial Intelligence, 17,* 17-45

Wolfson, S., & Landy, M. S. (1998). Edge- and region-based texture analysis mechanisms. *Vision Research, 38,* 439-446

Wyart, V., & Tallon-Baudry, C. (2008). Neural dissociation between visual awareness and spatial attention. *The Journal of Neuroscience, 28,* 2667-2679

Xiao, Y., Casti, A., Xiao, J., & Kaplan, E. (2007). Hue maps in primate striate cortex. *NeuroImage. 35,* 771-786

Yeari, M., & Goldsmith, M. (2010). Is object-based attention mandatory? Strategic control over mode of attention. *Journal of Experimental Psychology: Human Perception and Performance, 36,* 565-579

Zeki, S. (1973). Colour coding in rhesus monkey pre-striate cortex. *Brain Research, 53,* 422-427

Zeki, S. (1983). Colour-coding in the cerebral cortex: the reaction of cells in monkey visual cortex to wavelengths and colours. *Neuroscience, 9,* 741-765

Zeki, S. (1990). A century of cerebral achromatopsia. *Brain, 113,* 1721-1777

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

Zipser, D., & Andersen, R. (1988). A back-propagation programmed network that simulates response properties of a subset of posterior parietal neurons. *Nature, 331,* 679-684