Citation: Gheri, C. (2008). The role of salience on crowding and visual search in the context of synaesthesia. (Unpublished Doctoral thesis, City University London)

THE ROLE OF SALIENCE ON CROWDING AND VISUAL SEARCH IN THE CONTEXT OF SYNAESTHESIA

Carolina Gheri

Doctor of Philosophy

City University
Henry Wellcome Laboratories for Vision Sciences
Department of Optometry and Visual Science

December 2008

to my family

Contents

List of figures

Appendix - List of figures

Abstract

Visual ‘crowding’ is a phenomenon in which the identification of visual stimuli is impaired by nearby ‘distractors’. It occurs both for simple stimuli (oriented lines) and for more complex forms (letters). The literature on ‘crowding’ is reviewed, along with relevant literature on visual search and stimulus salience. Experiments are reported to test the idea that visually ‘salient’ stimuli can escape, in part, from crowding. The salience of stimuli was manipulated by varying their motion direction, colour or temporal frequency relative to distractors. Salience was also measured independently of crowding using the ‘pop out’ paradigm in visual search. Results showed that stimuli independently defined as salient did escape, in part, from crowding. A following experiment attempted to see whether the same would be true for the subjective colours experienced by synaesthetes. The hypothesis was that targets having a different synaesthetic colour from distractors would (a) ‘pop out’ in visual search and (b) escape from crowding. However, the expected ‘pop out’ effect did not occur, so the experiment was not extended to crowding. Finally, an experiment was performed to test the ‘compulsory averaging’ model of crowding versus the stimulus uncertainty (‘max rule’) alternative. Previous results supporting compulsory averaging were replicated, but the finding of a bimodal distribution of matches between target and a set of templates supported the ‘max rule’ model.

Acknowledgements

I would like to thank all the people who, during the last several years provided me the support and encouragement I needed.

Firstly I must thank Michael Morgan for having welcomed me inside his workgroup and for his precious supervision. His experience and love for the science were worth more to me than thousands of papers.

Josh Solomon for his gentle support and for the constant help he offered me.

Stefano Baldassi for giving me the opportunity of working in his lab in Florence and being able to collaborating with him.

Furthermore I have to really thank everyone I met at City, especially, Adrian, Brice, Valentina, Cristiano, Christophe, Franziska, Grace, Steve, Penny and Lucy. They have been with me during all this long experience and they shared with me both the difficulties and the beautiful moments I spent in London.

I would like to especially thank Marisa for her extraordinary friendship and for being there whenever I need her.

At last I have to thank Samuele, just for being the wonderful husband he is. I certainly could not finish this thesis without him next to me.

Declaration

I grant powers of discretion to the University Librarian to allow this thesis to be copied in whole or in part without further reference to myself. This permission covers only single copies made for study purposes, subject to normal conditions of acknowledgement.

Chapter 1

Introduction

1.1 Overview

In this thesis I will present and discuss my studies on visual perception, more specifically on the role of salience on crowding and visual search in the context of Synaesthesia. I will further discuss some modelling of crowding starting from the resulting evidence.

In this chapter I will introduce the main topics of the Thesis. It will be helpful to the reader to explain the actual historical progression of the topics studied, which differs from the order in which they will be described in the Chapters. The original intention of the work was to examine the claim that the subjective colours reported by Synaesthetes can aid them in early visual processing. I intended to use the techniques of visual search and of 'crowding' to examine this issue, and to devise experiments that would determine the stage of processing at which synaesthetic colours could influence performance measures. Unfortunately, a lengthy investigation of visual search in a sample of synaesthetes previously characterised in an fMRI study (2002), failed to confirm previous claims that synaesthetic colours influence visual search. It was therefore decided not to proceed with the intended study of 'crowding' in synaesthesia. Instead, it was decided to look at the role of 'salience' in general in normal subjects. I studied the definition and role of salience in visual search and crowding, and these are the investigations reported here. The experiment on synaesthesia has been relegated to the final Chapter. In addition, to studying the role of salience in visual search and crowding it was necessary to examine current theories of the set-size effect in search and crowding. This led to an experimental and theoretical investigation of the signal detection theory of the set size effect, which is reported in Chapter 4.

Finally, an experimental study of the averaging and 'max' rules in relation to crowding will be described in the Appendix. This study was carried out in collaboration with Stefano Baldassi in Firenze, and is in press in Vision Research.

1.1.1 Salience

In any moment of our life the brain processes and elaborates a continuous flow of information that arrives through our senses. To be able to do that a filter is used so that only a part of the information flow is really taken into account.

The visual stream, for example, can be filtered in many different ways. Sometimes we can decide a priori which is the most important information we have to pay attention to. If you are asked to check the text on the green label of a bottle on a shelf you will concentrate on the writing and ignore most of the other information about the product, for example you will not notice a damaged cap. This type of filtering is called a top-down process in which the filter we apply to the information stream depends on the knowledge of the object/target that has to be found or analysed.

At other times objects just attract our attention because they are particularly *salient*. Saliency can be defined as how much an object stands out from the background. Going back to the example above, you cannot avoid noticing a bottle with a red label while you are checking all the green labels. This is called a bottom-up process and it depends on the saliency of the object. In psychophysics salience is usually studied in the context of search experiments (see below).

1.1.2 Visual Search

Visual Search is a classic psychophysical paradigm in which the observer is asked to detect, and possibly to locate an item, the target, that is different from all the others present in the visual scene, the distractors. The number of distractors can be varied. The sum of all the items is the set size. Usually reaction times are measured. For a salient target the response time is faster than to a less salient target, by definition. In some cases, reaction times do not vary when more distractors are added to the visual display. This is the pop-out effect (Treisman, 1985, Wolfe, 1998).

But how can you decide if a target is salient or not and if one is more salient than another?

The target can differ from the distractor by a unique characteristic or feature: for example the observer has to detect a red line amongst green lines; this is usually called a simple feature search. It has been shown that in this case Reaction Times (RT) do not vary by adding more items to the display.

In other conditions, the target can have more than a characteristic that make it different from the distractors: for example the observer has to look for a horizontal green line amongst, vertical green lines and horizontal red ones. In this case the observer has to join two different features, colour and orientation, to detect the target. In this condition, called conjunction search, RTs are usually very sensitive to the number of distractors so that they are slower when more elements are added to the display: this is called the set size effect.

These findings and definitions represent the basis for one of the most important theories on the elaboration of the visual information: the Feature Integration Theory (FIT) (Treisman and Gelade, 1980). In this model the information can be processed either in a parallel way, as in the case of feature search, or in a serial way as in a conjunction task.

This FIT model has a number of known limitations, for example the asymmetry phenomenon. Given a target A, it can be easier (faster RT) to look for A amongst B’s than for B amongst A’s (Treisman and Gormican, 1988, Wolfe, 1998). This simple example gives evidence that measuring reaction time is not enough to define the salience of an object by itself (see Chapter 4).

Many computational models have been proposed trying to overcome the limitations of FIT theory. Several studies has been conducted to explain the mechanism of visual elaboration and giving salience a key role, but what really seems to be lacking is an objective measure of salience. The problem is that salience certainly effects the visibility and the detection of a target, but such a property is strongly related to the context in which the target is embedded and on the feature to which it is applied (Nothdurft, 2000a, Nothdurft, 2000b, Nothdurft, 2002).

1.1.3 Crowding

One factor affecting the salience of a target in peripheral vision is referred to as ‘Crowding’ defined as the impairment of the identification of a target in a cluttered scene (Stuart and Burian, 1962). In a classic crowding paradigm the stimulus is presented in the parofovea or in the periphery of the visual field. The observer has to keep fixation on the centre of the screen and has to identify the central target surrounded by other elements called flankers.

Crowding was first studied in the context of amblyopia (Ehlers, 1936) and in low vision patients in which central vision can be completely lost (Korte, 1923). Crowding has often been studied previously using letters as stimuli in the context of investigations of reading. Bouma (1970) suggested using simpler stimuli that could be controlled more easily, such as tilted lines. Such an approach permits a more precise investigation of the basic mechanism underlying the phenomenon of crowding. Quite a few models and ideas have been proposed, but none of them seems to completely explain all the effects that neighbouring items can have on the identification of the target. One of the initial aims of this work was to investigate whether the role of salience in crowding could help to solve some of the conceptual problems in this area.

1.1.4 Synaesthesia

Synaesthesia is a complex condition in which some people experience strong inter-modal perception. Examples include seeing colours and forms when listening to music or associating days of the week and months of the year with specific colours.

Synaesthesia has been reported historically by numerous artists who have attempted to describe their peculiar experience by painting, writing or composing music. Kandinsky has been claimed to be a synaesthete and Baudelaire and Rimbaud are well known for their poetry in which different senses are mixed together. Historically, Synaesthesia has been reported as more common in women than in men (Galton, 1880; Cytowic, 1989; Baron-Cohen, 1993).

Synaesthesia can be induced by using drugs as for example LSD (Hollister, 1968). There have been claims that Synaesthesia is just a normal period of brain development due to the interconnection between different cerebral areas that would generally disappear in older age (Maurer, 1993; Baron-Cohen et al, 1993), while others claim that Synaesthesia is a simple result of memory association created during childhood (Marks, 1975).

Recent studies have suggested that the colours experienced by synaesthetes can help them in visual search, perceptual grouping, and in overcoming crowding:

(1) If normal subjects try to find a ‘2’ amongst a set of ‘distracter’ digits such as 5’s their performance deteriorates as the number of distractors increases (the set-size effect). However, if the target is uniquely coloured it ‘pops out’ and performance shows little change with distracter number (Treisman and Gelade, 1980). Visual search with both real and synaesthetic colours was investigated by Palmieri, Blake, Marois, Flanery & Whetsell (2002) in a synaesthete WO, along with normal controls. WO responded more quickly than the mean of the controls, when the search item had a unique synaesthetic colour.

(2) A set of black 2’s forming a global shape such as a triangle does not, for normal subjects, stand out from a background of 5’s (Beck, 1982). However, if the 2’s are red and the 5’s are green, the global shape is seen much more easily. Ramachandran and Hubbard (2001b) investigated whether this was true for synaesthetic colours as well. They reported that two synaesthetes were significantly faster than 40 control subjects at reporting the shape formed by achromatic graphemes that had a different synaesthetic colour from the background graphemes. The experiment was later extended to six synaesthetes (Hubbard, 2005b) , five of whom showed superior performance to their control groups (n=20).

(3) Ramachandran & Hubbard (2001b, Ramachandran, 2001a) reported that synaesthetic colours could also reduce ‘crowding’. Graphemes presented in the periphery of vision are harder to identify when they are surrounded by other, flanking graphemes. Two synaesthetes were better than controls at identifying the target grapheme when it had a different synaesthetic colour. Subsequent research with six synaesthetes found the superior-to-normal effect in three of the six (Hubbard, 2005b).

As we shall see later, all this evidence is controversial.

1.2 Thesis Synopsis

The next Chapter (Chapter 2) will give a description of crowding. I will present a summary of different studies on the phenomenon and I will list several different effects that flankers can have on various aspects of of sensitivity and discrimination, by using different kinds of stimuli. In the last part of the chapter I will draw attention to some of the possible explanations given to describe the effect, in particular, the *Averaging model*, in which all signals are pooled together (Green and Swets, 1966) and the *Maximum-of-outputs rule* (Palmer et al., 1993) in which all signals are independently processed and compared so that the highest value is chosen. An experimental test of the two models of crowding will be described in a separate Appendix, so as not to disrupt the flow of narrative from crowding, through visual search, to synaesthesia.

Chapter 3 will give a review of Visual Search. It is very important to introduce this basic paradigm since most of the models discussed here in the context of crowding have been first tested for Visual Search tasks. Moreover this paradigm has been applied to an experimental study on Synaesthesia which will be described in Chapter 5.

The main aim of the studies reported in Chapter 5, was to investigate whether the subjective colours experienced by synaesthetes were equivalent to low-level sensory phenomena or if they were due to memory association.

Chapter 2

Review of crowding

The identification of a target in a visual cluttered scene can be very difficult especially when many elements are added to the display. Displacing the stimulus towards the periphery of the visual field and decreasing the target size or its distance from the surrounding elements makes the task even harder to carry out. This effect is called *crowding*.

2.1 General description

Spatial interactions are a critical feature of spatial vision that serve to sharpen perception of form and enable features to be grouped into forms. Spatial interactions may be inhibitory (reducing sensitivity) or facilitatory (enhancing sensitivity). Crowding is a form of inhibitory interaction in spatial vision that refers to the deleterious influence of adjacent flanks on visual discrimination.

Crowding has a practical significance in patients with low vision, due for example to age-related macular degeneration, one of the major causes of central vision loss. In these patients the central retina is damaged and they are forced to use the peripheral retina producing large deficits in normal everyday life tasks, such as reading. Slow reading in peripheral visual field is, in fact, mostly due to crowding among individual letters (Legge et al., 2001, Pelli, 2007); even when targets are scaled in size, the spatial extent and the intensity of the interaction are greater in peripheral than central vision (Legge et al., 2001).

Crowding has also been widely studied in normal observers. In traditional experiments the observer keeps his or her gaze on the fixation point and has to identify a target letter presented in the periphery. The target can appear either alone or surrounded by various number of letters, called flankers. In the condition when the flankers are present visual performance is worse than when the target is presented alone.

Early studies on crowding were carried out on letters and digits because these stimuli are widely used in various information displays in the real world. Crowding was discovered in the course of measuring letter acuity in patients with central field loss by Korte (1923) and in amblyopia by Ehlers (1936). Stuart and Burian (1962) coined the term “crowding” for the impairment of identification of a peripheral letter by neighbouring letters. A variety of other names have been used in the literature to describe the crowding effect, including: “contour interaction” (Flom et al., 1963), “interaction effects” (Bouma, 1970), “lateral inhibition” (Townsend et al., 1971), “lateral masking” (Wolford, 1983, Taylor, 1972, Geiger, 1986), and “masking” (Anstis, 1974).

The effect has been also measured with different stimuli in different tasks such as two-bar resolution (Takahashi, 1967), Vernier discrimination (Westheimer and

Hauske, 1975, Levi et al., 1985, Levi and Klein, 1985), stereopsis (Butler and Westheimer, 1978) and line orientation sensitivity (Westheimer et al., 1976).

In foveal vision results showed that crowding only occurs over very small distances (Toet and Levi, 1992, Flom et al., 1963, Liu, 2000), or it does not occur at all (Strasburger et al., 1991). In peripheral vision, crowding occurs over distances several times larger than the target size. The mechanism for crowding is still not completely understood. The remainder of this Chapter reviews the theories and evidence.

2.2 What is the effect of flankers?

2.2.1 Letter Identification

2.2.1.1 Cortical Magnification

Historically, visual acuity was defined in terms of the smallest stimulus that can be identified with a specified level of accuracy. Aubert and Foerster (1857) showed that peripheral vision has a lower acuity than foveal vision, and that the decrease in visual acuity can be compensated for by increasing the stimulus size. More recently, visual acuity has also been measured in terms of contrast sensitivity, and the interaction between stimulus size and contrast is important in measuring the fall off in visual performance with retinal eccentricity.

The map of the visual field on the visual cortex is non-linear (Virsu and Rovamo, 1979, Cowey and Rolls, 1974 , Daniel and Whitteridge, 1961 , Van Essen et al., 1984); which means that changing the size of a stimulus changes the amount of cortical tissue stimulated, depending on the location of the target in the visual field. A simple scaling hypothesis suggests that, in order to have performance that is invariant with eccentricity, we have to be sure that the print size is appropriately scaled.

In general, contrast sensitivity in the periphery is severely degraded for small letters, for example, six times more contrast is required to recognize 3-mm targets compared with 10-mm ones (Strasburger et al., 1991). If letters are big enough (above about 10 mm) performance is independent of size. Thus, the function relating contrast threshold to target size has a shape similar to that one in the fovea, but it is shifted, to the right and upwards, towards larger target sizes (Pelli, 1999).

2.2.1.2 The effects of flanks on threshold size and contrast threshold for letter identification

Readability of character groups cannot easily be deduced from individual character recognition (Townsend et al., 1971). Specifically, as already mentioned, the recognition of a target character is more difficult in the presence of neighbouring characters. Crowding was in fact discovered measuring letter acuity in patients with central field loss by Korte (1923) and in patients with amblyopia by Ehlers (1936) and Woodrow (1938).

Strasburger (1991) studied the recognition of numeric characters in foveal and eccentric vision by determining the contrast required for 67% correct identification. Groups of three digits were presented, and the subjects were asked to identify the middle one. The centre-to-centre distance from the flanking digits to the target digit was set at two letter-size units, which creates one blank space between the digits. As a consequence, the larger target sizes had larger distances to their flanking digits. It was found that target character size and distance of flanking characters were the most significant variables. The influence of these parameters was different in the fovea and in the periphery, validating previous qualitative studies (Aubert, 1857). In the fovea, contrast thresholds were mainly independent of target size when it was above a certain threshold (about 0.6°). For target sizes below this limit, contrast thresholds increased so much that, for the smallest size used, the maximum contrast attainable was insufficient for target recognition. In peripheral vision, where sensitivity is systematically reduced, the effect was similar to that one observed in the fovea, but it was shifted toward larger target sizes. Concerning the flankers distance, in the fovea contrast thresholds were found to be independent of character separation. For peripheral viewing, a large crowding effect was measured below a certain character distance, such an effect vanished at larger separation values. These results also confirmed Bouma’s law (1970), which states that critical spacing is proportional to eccentricity independently of the letter size. Additionally it was shown that the effect of flankers is small, or possibly absent, in the fovea, but that already at 2 degrees from the fovea it is quite pronounced. The same author showed that the effect occurs only for letters smaller than a certain critical size and that it depends on eccentricity (Strasburger et al., 1991).

Crowding has often been compared to masking. Masking usually refers to the effects of spatially-overlapping masks on target identification or detection. Masking and crowding have sometimes been confused since they both impair visual discrimination (see section 2.2.3.4). Chung et al (2001) aimed to compare the two phenomena in different experiments in order to highlight the spatial properties of crowding. They first measured contrast thresholds for identifying the middle target letters in trigrams for a range of letter spacing (from 0° to 5° eccentricity) while the flankers contrast was kept constant. Letters were digitally filtered using a set of bandpass filters. They found that crowding exhibits spatial-tuning functions like masking, but with generally broader bandwidths than those for masking. In the second experiment, they measured the contrast thresholds for identifying the middle target letters in trigrams for a range of flanking letter contrasts at 5° eccentricity. Thresholds for identifying the middle target were not affected by low contrast flankers while at high flanker contrast, the thresholds increased with contrast.

In the final experiment, they varied the contrast ratio between the flanking letters and the target letters. The aim of this experiment was to see if crowding would be highest when flankers and target had the same contrast. This prediction was based on previous studies showing that the crowding effect was stronger when flankers and target share the same properties (Andriessen and Bouma, 1976, Kooi et al., 1994, Nazir, 1992) (see section 2.2.2.4). The magnitude of crowding actually increased monotonically with contrast ratio, showing that contrast difference is not the important variable.

Pelli et al (2004) presented a similar study in which they compared crowding and masking. They measured contrast thresholds for identifying a letter in the middle of a trigram. They varied spacing in between letters, eccentricity, and the size of both target and flankers. Considering the spacing and the eccentricity issue they confirmed the Bouma’s rule (1970). Regarding target size, their results agree with a previous study (Strasburger et al., 1991), in that critical space does not vary with target size while threshold elevation increases as a function of target size, since contrast sensitivity for letters depends on size (Pelli, 1999). The size of flankers does not seem to have a great effect on crowding, but appears to depend more on eccentricity and spacing.

Pelli et al (2004) also measured letter identification thresholds in the periphery varying the number of the flankers. Their results show that threshold elevation increased when flankers are increased from 1 to 2, but that thresholds do not vary much more if the number of flankers are increased from 2 to 4. These results are also confirmed by other studies not using letters (Wilson et al., 1997, Toet and Levi, 1992). Moreover Pelli et al (2004) varied the contrast of the flankers. They again measured contrast threshold to identify a middle letter in the trigram. They found what they called ‘mask threshold’, which is the level at which thresholds start to increase, and a ‘mask saturation’, at which contrast thresholds increase no further. The interesting finding was that these critical mask levels are not spacing dependant.

Finally they also investigated whether the shape or the complexity of the font could affect crowding, but threshold elevation in letter identification did not vary depending on the font of the letters.

2.2.1.3 Reading and Crowding

A more ecological context in which print size can have an important role in letter identification is *reading*. The idea that letter size limits reading is ancient. The classic reading rate curves are all plotted as a function of letter size. In general as text size increases, reading rate rises abruptly from zero to maximum rate. This classic reading-rate curve consists in fact of a cliff and a plateau, characterized by two parameters, *critical print size* and *maximum reading rate* (Legge, 1985); this basic result is well established but as yet unexplained.

According to the scaling hypothesis (Flom et al., 1963, Levi et al., 1985, Chung, 1998), reading performance in peripheral vision is the same as in central vision, except for a scaling factor in print size. Moreover the scaling hypothesis predicts that the reading speed versus print size plot is shape-invariant in peripheral vision and that the maximum reading speeds in central and peripheral vision are identical if letters are properly scaled (Latham, 1996a). Chung et al. (1998) evaluated the effect of print size on reading speed at different eccentricities in normal peripheral vision. Six normal observers read aloud single sentences (11 words in length) presented on a computer monitor, one word at a time, using rapid serial visual presentation (RSVP:Potter, 1984). They measured reading speeds for eight print sizes at each of six retinal eccentricities, from 0 (foveal) to 20 deg. Contrary to the simple scaling hypothesis, it was found that the maximum reading speed is not invariant with eccentricity, even when the print size was scaled, and it was concluded that print size is not the only factor limiting maximum reading speed in normal peripheral vision. The same authors obtained evidence that the visual span, the number of characters that can be recognised in a glance (Legge et al, 1997) reduces in size in the periphery (Legge et al., 1983, Legge et al., 2001). Such a reduction has been attributed to poorer spatial resolution caused by greater convergence of photoreceptors on to retinal ganglion cells (Virsu and Rovamo, 1979). The target was controlled in following studies to be equally detectable by either type of photoreceptor. The authors actually used a reading paradigm in which the text could only be seen by rods. Reading rates were measured with the text presented in different parts of the visual field. Observers were always faster reading the text when it was displayed in close proximity to the fovea. This actually proved that the intrinsic differences between cones and rods is unlikely to be a factor accounting for the slow reading speed in peripheral vision (Chaparro, 1993).

Pelli et al (2007) tried to give a complete explanation of visual span and of the reading rate curve by including crowding. They established a strong link between reading and crowding and showed that “visual span” is the “uncrowded span”, which is the number of characters that are not crowded. Finally, they demonstrated that only crowding determines the positions of the cliff (critical print size) and the plateau (maximum reading rate) of the reading rate curve. Their results suggested that crowding, and not acuity, determines the position of the cliff and that therefore performance of the flanked letter identification task depends on spacing, not size. Levi et al. (2007) confirmed this finding showing that doubling the normal letter spacing in the text shifts the reading rate curve, plotted as a function of size. This shows that letter spacing is important, even when size is controlled.

2.2.2 Orientation Acuity

Despite their practical significance, letters are poor candidates as visual stimuli for crowding. How these complex shapes are processed and encoded is not well understood and this makes it hard to analyse the underlying mechanisms of crowding when letter identification is employed. Following an influential paper by Bouma (1970) there has been a move towards using simpler stimuli such as lines, gratings or Gabor patches.

2.2.2.1 *Effects of flanks on orientation acuity*

Studies of orientation identification of lines or Gabor patches have shown similar crowding effects to those using letters. (Korte, 1923, Stuart and Burian, 1962, Flom et al., 1963, Bouma, 1970). The extent of crowding varies with eccentricity, the distance from the fixation point, and with critical spacing i.e., the least distance target-flanker at which threshold elevation disappeared.

The crowding range is roughly half the target distance from fixation (Bouma, 1970, Strasburger et al., 1991, Toet and Levi, 1992). This result, that has been named Bouma's rule, has been confirmed by other studies (Pelli et al., 2003), and it has been found to be independent of the target size (Strasburger et al., 1991, Tripathy and Cavanagh, 2002). One advantage of using gratings or Gabor patches as stimuli is that their spatial frequency content can be controlled separately from their overall size.

2.2.2.2 *The effects of flanks on orientation biases*

The discussion of flanks has so far concerned only effects on acuity. However, another effect of flankers appears in the case when the central target seems to be tilted away from its true direction. Such an effect is called 'tilt illusion' (Gibson and Radner, 1937, Blakemore et al., 1970, Over et al., 1972) and can take the direction of *assimilation*, the tendency to perceive vertical targets as tilted in the same direction as slightly-tilted distractors, or of *repulsion*, the tendency to perceive vertical targets as tilted away from more-oblique distractors. Solomon et al (2004) used a configuration of three Gabor patches, that could have different positions on the display, identified as 3, 6, 9, 12 'o clock, presented for 100 msec. The task for the observer was to judge if the central Gabor was tilted CW (clockwise) or CCW (counter clockwise) from the vertical. Both biases and sensitivities caused by flanks were measured at an eccentricity of 3.7° . When the tilt of the flanks relative to the target were small, biases tended to have the same sign (assimilation), this effect was quite small compared with the large, opposite-signed biases (repulsion) induced by grossly oblique (i.e. ±22.5° and ±45°) distractors. Vertical and horizontal distractors produced smaller biases than oblique flankers but in some cases their effects were significant. Oblique distractors responsible for the largest biases were also those responsible for the largest drops in sensitivity even though vertical distractors also produced sizeable losses of sensitivity. Horizontal and nearly horizontal distractors produced the smallest losses of sensitivity, but even these were significant in most cases. If the distractors were tilted in opposite directions, the biases were small but the sensitivities were similar to those measured when distractors had the same tilt.

Li et al. (2000) compared the effect of contextual stimuli on human orientation discrimination and on responses of neurons in V1 of monkeys. The stimulus was a line that was presented in various contextual conditions. Their data showed that contextual interference observed in human orientation discrimination can be related to contextual inhibition of neuronal activity in V1. Generally when human performance was impaired, for most of V1 neurons, responses elicited by the target line were suppressed by the contextual patterns. If the target was surrounded by orthogonal lines psychophysical discrimination was slightly better. Thresholds were elevated for randomly oriented lines and even more for lines parallel to the target. Correspondingly, the firing of V1 neurones to the target line were suppressed. Straight lines gave the strongest interference with human orientation discrimination and produced the strongest suppression of neuronal responses. The human psychophysical data and the single neurone recording were not always in agreement. Elevation of human orientation discrimination thresholds was highest when a mask covers only the immediate vicinity of the target line and increasing the masking area resulted in less interference. On the contrary, suppression of neuronal responses in V1 increased as the masking size became larger. The authors proposed that these results suggested a figure-ground segregation process that is not located in V1.

2.2.2.3 The effects of flanks on threshold contrast for orientation classification

Pelli, Palomares, and Majaj (2004) measured contrast thresholds for identifying the tilt of gratings. Signal and flanker were randomly tilted ±45° on each trial and spacing in between them was varied. In the detection task the observer was required to choose which of two intervals in a 2AFCT task contained the signal grating. In the identification task there was only one interval and the observer was asked to identify the orientation of the target. Thresholds were elevated only for narrow spacing (smaller than 1°-2°) in both kinds of task.

Other studies investigated the role of difference in contrast between target and flankers (Felisberti et al., 2005). The display consisted of a central target sounded by 4 flankers. Either the target contrast was kept constant and the flankers' varied or the other way round. The array was flashed for 100 ms at 3.8 deg eccentricity. The tilt was sampled from 5 possible tilts; each block consisted of 100 trials, 20 for each tilt. The cue was present in some cases, and the location changed so that the target was equally presented in each position. No post-stimulus masks were used. The results showed that contrast threshold elevation was reduced by increasing the contrast of the target relative to the distractors, but targets of smaller contrast than the distractors were even harder to identify than those of the same contrast.

2.2.2.4 Other known effects of flanks on orientation classification

Wilkinson et al. (1997) reported that reducing the number of flanking gratings from 14 down to 2 did not significantly reduce their effect on the discriminability of the signal. Similarly Toet and Levi (1992) measured the effect on orientation discrimination of a 'T' target of one or two collinear 'T' flankers (composed by two lines of the same length) which were randomly oriented up or down with the effect of one flanker. They made measurements for three different eccentricities (0°, 2.5° and 10°). A threshold elevation was registered only when two flankers surrounded the target.

On the other hand, Poder (2006) demonstrated that the crowding effect was remarkably reduced by increasing the number of distractors (up to 168 elements). Similar results were observed when the target and distractors were of the same colour and when only a differently coloured circle indicated the target location. Poder suggests that the results can be explained by bottom-up ‘‘attention’’ that facilitates the processing of information from salient locations in the visual field (Poder, 2006). We shall consider this notion of bottom-up salience in a later section (see Chapter 3).

Livne & Sagi (2007) investigated the effect of flankers' configuration on crowding. They compared a few different configurations (smooth contour versus random) for eight Gabor patches surrounding a central one. Orientation discrimination and contrast detection of the central Gabor were measured. These measurements revealed differences in the magnitude of the effect of crowding for different configurations, especially on the discrimination task. The crowding effect was stronger when random configurations were used and was reduced considerably when a smooth one was used.

In another study, Malania et al (2007) reported that interference caused by flankers in Vernier offset discrimination depends on their proximity. Both the length and the pattern of the flankers would influence this effect. Except for the equal-length condition in almost all configurations, the threshold elevation is reduced when the flankers are part of a more extended configuration.

Crowding shows a well-established spatial anisotropy, such that more peripheral flankers than the target have a greater effect than those nearer the fixation point (Shaw, 1969; (Bouma, 1970, Bouma, 1973, Chastain, 1983, Shaw, 1969). This result has been recently confirmed using Gabor patches as stimuli by Petrov and Popple (2007). In their experiment they asked the observers to identify the tilt of three horizontally aligned Gabors with 45° left or right from the vertical, shown to the left of fixation for 150 ms at 6° eccentricity. Next, a column with eight possible choices was shown to the right of fixation and the observer had to indicate the matching triplet. They found that the outmost target was much less crowded in comparison to the other two.

Another aspect that has been investigated is the target-flanker similarity. Kooi et al (1994) measured the effect of colour difference on crowding. Two out of six observers suffered crowding even when the target had a unique colour. The authors suggested that these two observers might not have experienced an enough strong perceptual difference between the target and the flankers. Following these results, Ramachandran & Hubbard (Ramachandran and Hubbard, 2001b, Ramachandran, 2001a) studied crowding in the special condition of colour-grapheme Synaesthesia. They found that two synaesthetes were better than controls at identifying the target grapheme when it had a different synaesthetic colour from the flanks. The same effect was also confirmed by (Hubbard, 2005b) in three of the six observers. These results are quite interesting because, if they can be confirmed, it would suggest that synaesthetic induced colours act as real ones. We will further discuss these results in the Chapter 5.

Felisberti et al. (2005) described a series of experiments manipulating target salience. A cue indicating the position of the target improved performance when the position was otherwise unknown and if the spatial separation between target and distractors was large, but the cue had no significant effect when separation was small. The ineffectiveness of the cue at small separations is evidence for crowding. Felisberti et al. (2005) also investigated differences in depth and showed that putting the target and distractors in different stereoscopic depth planes decreased crowding.

The relation between crowding and salience will be discussed in greater detail in a later section (see Chapter 3).

2.2.3 Contrast Detection

2.2.3.1 The effects of flanks on contrast detection

In a detection task, the target is randomly presented in one of two consecutive intervals. The flankers are displayed in both intervals, independently and randomly selected for each interval. Observers have to report in which interval the signal is present.

Polat and Sagi (1993) measured contrast detection thresholds for a foveal Gabor signal flanked by two high contrast Gabor patches (‘flankers’) as a function of their distance from the target. Their data showed a suppressive region extending to a radius of two wavelengths, in which the presence of high contrast masking signals had the effect of increasing target threshold. Beyond that, they describe a much larger facilitatory region where contrast detection threshold decreased again. They also showed that detection of a Gabor patch could be facilitated when flanked by low-contrast Gabors of similar orientation. Such a detection facilitation was maximal for flanking stimuli that were spatially proximal and oriented collinearly with the target (see also Solomon et al., 1999). Detection facilitation decreases as the target and flankers vary from collinear (Kapadia, 1995, Williams and Hess, 1998), although others have demonstrated significant target facilitation also occurs when flankers are oriented orthogonally to the target (Cass, 2005, Yu, 2002).

This foveal data has been compared with studies in the periphery. Williams and Hess (1998) measured contrast detection thresholds for a central patch with and without flanking patterns. In the fovea, the presence of flankers increased the sensitivity to the central patch, in agreement with the results of Polat and Sagi (1993). In the periphery, however, they found no facilitation.

Their results were confirmed by Levi and Klein (2002a). They considered whether crowding in peripheral vision can be explained on the basis of contrast masking by flanks. To test this hypothesis, they measured and compared crowding in a direction of identification experiment with masking by flanks in a detection experiment. They examined the relationship between the suppressive and facilitatory interactions in peripheral contrast detection and crowding. The results show again that unlike the foveal results, peripheral crowding shows separate characteristics. Their results suggest that inhibitory spatial interactions in peripheral crowding extend over larger distances than in the fovea for targets of the same size (Levi et al., 2002a). It was found that the critical distance for crowding is approximately 0.1 times the target eccentricity. In general, the results show that observers can easily detect the features that compose the signal targets, in this case Gabor patches, even if crowding was very strong. The authors suggest that peripheral crowding happens because the target and flanks are combined or pooled at a second stage and that in peripheral vision; this pooling takes place over a larger distance. This suggestion was also made by Pelli, (2004).

2.2.3.2 The effects of surrounds on contrast discrimination

Snowden and Hammett (1998) investigated perceived contrast, contrast detection and contrast discrimination of a central grating pattern as a function of surround contrast both in fovea and periphery (5 deg and 10 deg eccentricity). In the fovea, surrounds had an effect only at low test contrasts. There was no effect on detection thresholds, and only a small effect on discrimination at low pedestal contrasts. In the periphery the effect of the surround was generally greater and affected a larger range of test contrasts. Detection thresholds were increased by the surround. The contrast discrimination dipper function (Foley, 1994), was shifted upwards and rightwards, as in previous studies of adaptation and masking by Ross et al. (1993). Snowden and Hammett argue that the effects of a surround are essentially the same as that of a traditional spatially coextensive mask, as modelled by Foley. The significance of this study for the theory of crowding is that it suggests that there may be a role for masking in crowding, although as we have seen earlier, there are important differences between crowding and masking in other respects.

2.2.3.3 Other effects of flanks on contrast

Pelli et al. (2004) measured detection thresholds for a letter among letters as a function of flanker spacing. Contrast threshold elevation for the detection task was smaller than for the identification task and it extended over almost the same spatial range. Pelli et al. also assessed the effect of eccentricity and size on critical spacing: they measured detection thresholds for 0.75° Sloan letters for three different eccentricities (2°, 4°, and 8° in the right visual field). Results show that the critical spacing for detection is independent of eccentricity. They performed the same test for three letter sizes, 0.75, 1.5, and 3 deg, at 8 deg in the right visual field. The results showed that the critical spacing for letter detection is proportional to size. In summary, the authors showed that the critical spacing for letter detection was proportional to size while letter identification is independent.

These results confirmed previous data by other authors using different stimuli. Levi et al. (2002a) reported that observers could easily detect the features that compose the targets (Gabor patches) under conditions where crowding was strong. Threshold elevation for detection was much weaker and less extensive than threshold elevation for the identification task.

2.2.4 Crowding and masking

The terms ‘masking’ and ‘crowding’ have not been consistently defined in the literature. The term *crowding* is ordinarily used to describe the fact that adjacent objects (letters or flanks) reduce the discriminability of a target. Similarly, the term *masking* is usually used to describe the fact that a pattern can reduce the discriminability of a target when pattern and mask are spatially overlapping (Legge and Foley, 1980, Foley, 1994). However, when nearby but non-overlapping stimuli reduce the detectability of a simple target like a Gabor patch, the effect has often been called ‘masking’ (Pelli et al., 2004, Polat, 1993, Shani, 2005). Since crowding and masking both impair visual discrimination it is reasonable to ask whether they are two sides of the same coin and if they possibly share a common mechanism. Studies that have addressed this question (Chung et al., 2001, Pelli, 2004), have revealed clear differences between the stimulus parameters determining masking and crowding. In masking, critical spacing is proportional to signal size and it is independent of eccentricity (Pelli et al., 2004, Shani, 2005, Polat, 1993). This effect can be either positive (reduced contrast detection threshold) or negative (increased contrast detection threshold), depending on the target–flanker distance and the stimulus configuration (Polat, 1993). In crowding, on the other hand, the critical spacing is proportional to eccentricity but independent of size and, above all, it is always negative.

2.3 Possible explanations for crowding

Although crowding has been studied psychophysically since 1923 (Korte, 1923, Stuart and Burian, 1962, Flom et al., 1963, Bouma, 1970) its physiological basis is still not well understood. Here I will present some of the main theories that have been proposed.

2.3.1 Spatial resolution hypothesis

In their classical study, Flom, Weymouth, and Kahneman (1963) estimated the extent of crowding as a function of eccentricity by having observers judge the orientation of a near acuity threshold Landolt C while varying the distance of surrounding flanks from the ‘C’. Flom et al. suggested that the distance over which spatial interaction occurs is related to the size of the receptive fields that are most sensitive to the target. Because peripheral vision is characterized by reduced visual acuity, larger receptive fields will be engaged (because small receptive fields have low sensitivity), and this scale shift will result in proportionally larger crowding distances. This scale shift hypothesis makes several predictions. First, it predicts that in peripheral vision, the spatial extent of crowding (with broadband stimuli) will scale with (be proportional to) the uncrowded acuity. For letter acuity (Flom et al., 1963) and Vernier acuity (Levi et al., 1985, Levi and Klein, 1985), the spatial extent of crowding appears to scale with the unflanked letter or Vernier acuity in amblyopia and peripheral vision, respectively. Second, the scale shift hypothesis predicts that the spatial extent of crowding will depend on the size or spatial frequency of the target.

Most of the previous studies of peripheral crowding use broadband stimuli (e.g., letters) that are close to the acuity limit. Thus, crowding in peripheral vision is typically measured with larger (broadband) stimuli than in the fovea, and the resulting data may indeed reflect a shift in spatial scale. Levi et al (2002a) used a spatially band-pass filtered E stimuli composed of narrow-band features and showed that peripheral crowding extends over a greater distance even when tested with the same size (and spatial frequency) stimuli as the foveal stimuli. Thus the authors argued that it is unlikely that the extended peripheral crowding reflects a simple shift in spatial scale of first stage filters toward larger (low spatial frequency) filters at an early stage of processing in V1.

2.3.2 Lateral Inhibition (Masking) hypothesis

Crowding has also been often used as synonym of the broader term ‘lateral masking’ (Townsend et al., 1971, Wolford and Chambers, 1984, Mansfield, 1998). Lateral masking refers to any effect on the detectability, discriminability or recognition of a target by non-overlapping spatially adjacent patterns. The use of the term ‘lateral masking’ to refer to crowding leads to the implication that crowding can be a form of masking in which the visibility of a target is affected by a spatially superimposed masker. Therefore, as we have seen in some sections above, the goal of many different studies has been to compare several properties of crowding with those of masking, in order to determine if they share the same effects and, presumably some common underlying mechanisms (Levi et al., 2002b, Pelli et al., 2003).

For a long time the two effects have been confused and only lately some differences have been found. Several approaches agree in concluding that crowding and masking are likely to share the same first stage filtering process, and perhaps have similar second stage feature integration (Pelli et al., 2003, Chung et al., 2001).

Surround suppression and crowding have many similarities, and surround suppression could, in principle, cause crowding by eliminating weak but important features of the target (e.g., the difference between letters G and C). The lateral inhibition model was inspired by an observation that similar letters produce stronger crowding (Estes, 1974, Estes, 1972). Using the observation that confusable (similar) letters produce stronger crowding, Estes suggested that crowding happens because similar features of nearby letters inhibit each other.

This class of explanations cannot explain several recent results. For example it does not explain the strong inward–outward anisotropy, in which there is greater crowding of stimuli nearer to the fixation point (Chastain, 1986, Petrov, 2007). Moreover, following this theory, we could assume that the high-contrast Gabor patches could suppress one another to such a point that the perceived contrast of the Gabor patches would be dramatically reduced. In fact, no such effect was observed by (Pelli et al., 2004, Petrov, 2007) who found that identification WAS impaired, when the target was still visible.

Polat and Sagi (1994) have argued that the facilitation by remote flankers observed in foveal detection is a consequence of excitatory long-range horizontal connections between neurons with like orientations in the cortical area V1. These horizontal connections may be excitatory and inhibitory (Fitzpatrick, 2000, Gilbert, 1998), and extend up to about 1 to 2 mm in the primate area V1 (Rockland, 1983, Blasdel, 1985, Lund, 1993, Amir, 1983). The long-range inhibitory connections have approximately the length to account for peripheral crowding. On the other hand, this close correspondence may be a coincidence, because long-range connections appear to be too short to account for foveal crowding (Levi et al., 2002b). Moreover, the fixed cortical distance of long-range connections predicts interactions over a fixed retinal distance, rather than interactions that are related to target size in the fovea. In addition, it is unclear why long-range interactions in masking should be different from long-range interactions in crowding.

2.3.3 Spatial averaging hypothesis

Parkes et al. (2001) suggested that crowding is related to texture perception. According to their compulsory averaging model, in crowded arrays, target and flankers are processed independently and their individual orientations are computed but only their spatial average is available for reporting. Their evidence came from experiments in which observers had to report the direction of tilt (clockwise or anticlockwise or vertical) of a central Gabor patch surrounded by a variable number of flanks. When the flanks were all vertically oriented, tilt thresholds increased linearly with the number of flanks, as predicted from linear averaging. When surround patches were tilted in the same direction as the target, thresholds fell linearly with the number of tilted flanks, again as predicted from averaging. Finally, when some of the patches were tilted in the opposite direction from the target, but by a smaller absolute amount, thresholds were well predicted from averaging, echoing a previous result of Morgan, Ward and Castet (1998).

The averaging hypothesis has been challenged in recent papers by Baldassi et al. (2006) who claimed that spatial uncertainty (specifically, the ‘Max’ rule of signal detection theory) gives a better account of the results. We shall examine this issue experimentally in Chapter 4.

Similar to the suggestion of averaging, Levi et al. (2002a) and Pelli et al. (2004) argued that signal processing happens in two stages in the periphery. At the first stage, individual features are detected and surround suppression occurs. At the second stage, the feature information is pooled over an area of visual field proportional to the stimulus eccentricity, which results in crowding. To support this hypothesis, Levi et al. show that observers could easily detect the targets under conditions where crowding is strong. He et al. (1996) used adaptation to grating stimuli to show that detection and crowding are happening at different processing stages.

Pelli et al., (2004) also suggested that crowding occurs when subjects have to identify a complex feature in the stimuli. To do so, they have to integrate the outcome of several basic feature detectors. This integration is suggested to be carried out within an integration field, which operates over increasingly larger areas as one moves further into the visual periphery. As a result from this integration fields’ size limitation, in crowded displays, even when centred on the target, the integration process incorporates information also from flankers’ signals. Thus, according to this explanation, the extent of crowding is determined exclusively by target–flanker separation and its ratio with eccentricity. A consideration on configural effects strongly suggests that such an explanation may at least be incomplete. If crowding was the consequence of only such hard-wired spatial mechanisms, then no differences would be expected from the different configurations of the stimulus used, but this prediction seems not to be confirmed by experimental results (Felisberti et al., 2005, Livne and Sagi, 2007).

2.3.4 Attentional hypothesis

Intrilligator and Cavanagh (2001) have suggested an attentional resolution model of crowding. According to this explanation crowding represents the limits of attentional acuity, which is quite coarse in the periphery. When items are spaced more closely than the smallest possible region of attentional selection, more than one item will be selected as a group and access to the individual identities is not possible. If the target is the only item within the selection region, then it can be identified until it is too small to be visually resolved.

Yeshurun & Carrasco showed similar results (1998, 1999, 2000) without using a crowding paradigm. They used a peripheral precueing to explore the effect of covert transient attention on performance in spatial resolution tasks in the periphery, where the effect of crowding is usually found. They found that directing attention to the target location enhances performance where the resolution is too low . The authors suggest that attending to a location is similar to reducing the size of the corresponding filters.

2.3.5 Uncertainty hypothesis

Crowding can also be the result of positional uncertainty (Huckauf, 2002, Strasburger, 2005). According to this type of explanation, spatial localization is limited and thus crowding is the result of reporting the properties of a flanker instead of the target. Unlike in the feature integration account where features are free-floating, here complete objects rather than features are assigned wrong locations. This theory is supported by results showing that in letter crowding experiments, subjects incorrectly report a flanker letter present in the display more often than a letter not present (Strasburger, 2005, Huckauf, 2002). However, this result would also be predicted by selection of features from flanking letters and their inappropriate integration into the letter-identification stage.

2.4 Conclusion

This Chapter has shown that there is a wide variety of theoretical explanations of crowding, and that no theory so far explains all the data. Loss of feature binding, averaging, and failure of attentional resolution may all be involved. An important issue, identified in section 2.2.2.4 is whether crowding is affected by target salience (‘pop out’). Experiments have shown that making targets different from distractors along certain dimensions, such as colour, can reduce crowding, but there has been no quantitative investigation of the relationship between relief from crowding and ‘pop out’. Such an investigation would be possible if we were to measure ‘pop-out’ independently from crowding, using the set-size effect in visual search. This is the aim of the investigations reported in the next two Chapters. Before turning to the experimental chapters we review the literature on ‘pop out’ and visual search.

Chapter 3

Review of Visual Search

In everyday life our brain has to organise a lot of visual information: for example, looking for something or somebody in a complex and packed environment. This issue is studied in a the psychophysics laboratory through Visual Search tasks.

3.1 General description

3.1.1 The classic paradigm

In the classic visual search paradigm (Treisman and Gelade, 1980, Treisman and Souther, 1985) the observer has to decide whether a pre-specified target appears among some number of *distractor* items. The total number of elements in the display is known as the *set size*. The schematic representation of these components is shown in Figure 3-1.

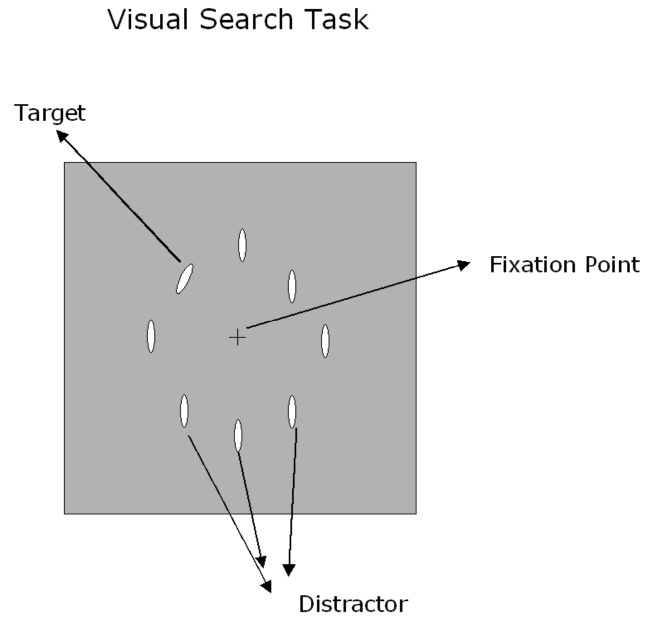


Fig. 3-1: Schematic representation of a visual search task and of its main components. In this case, the target differs in orientation

In some percentage of the trials, usually 50%, the target is present. In the other trials only distractors are present. Average reaction time is usually plotted as a function of set size. The slopes of these curves are used to quantify performance. The slope for target-present trials often differs from the slope for target-absent trials.

During the 70s and the 80s most of the models of visual search had both parallel and serial architectures. When set size did not affect performance, the search was assumed to be mediated by pre-attentive, unlimited capacity and parallel processes. When set size did affect performance, the search was attributed to attentive, limited-capacity and serial processes (Neisser, 1967, Kinchla, 1974, Shiffrin, 1977, Schneider, 1977, Sternberg, 1969).

3.1.1.2 FIT theory : Feature search and conjunction search

The notion of a division between parallel and serial search was central to the Feature Integration Theory (FIT), first proposed by Treisman and Gelade (1980). FIT was particularly influential because it offered an explanation for the difference between feature search, where the target differs from the distractors along only one stimulus dimension like orientation (Foster and Ward, 1991b), colour (D'Zmura, 1991, Carter, 1982) or curvature (Treisman and Gormican, 1988), and conjunction search, where the target differs from some distractors along one dimension and differs from others along another dimension (Treisman and Gelade, 1980).

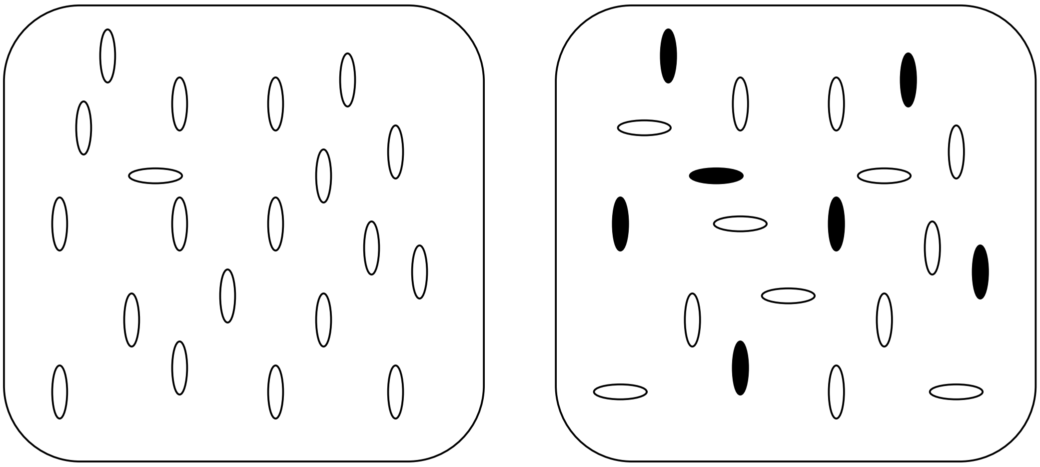


Fig. 3-2: Example of search displays in which the target is defined by a single feature (left panel), in this case horizontal orientation, or by a conjunction of features in a unique combination (right panel), in this case black polarity and horizontal orientation.

There is typically little-to-no effect of set size in feature search. The target is said to ‘pop-out’ from the display. On the other hand, Treisman and Gelade (1980) found that conjunction searches produced linear reaction-time-versus-set-size functions with slopes in the region of 40-60 ms per item in target-absent trials and 20-30 ms in target-present trials. According to FIT, features are analysed independently, at a pre-attentive stage and in a parallel way. The result of this analysis is a family of spatially (or retinotopically) organised maps; one for colour content, another for orientation content, etc.

In a conjunction search, no single feature map carries information about the uniqueness of the target. Different features at a single location would have to be *integrated* first. This, according to FIT, happens when visual spatial attention is focused on a particular region of interest. Due to the limited capacity of attentional resources, separate regions of interest would be attended in sequence, one after another.

Serial searches like this were thought to be self-terminating. The slopes for target-present searches were usually roughly half those for target-absent searches (1:2 ratio). In the former case, on average, attention would need to process only half of the items before the target would be found. In the target absent condition, attention would need to pass through all items before observers could be confident that no target was present.

All subsequent theories of visual search can be considered either reactions to, or modifications of FIT. Most are concerned with one or more of FIT's various limitations (described below). Most importantly, these newer theories make quantitative predictions and incorporate our greatly improved understanding of neuronal activity.

3.1.1.3 Limits of FIT theory

The serial-parallel dichotomy has been questioned by many authors (Neisser, 1963, Treisman and Sato, 1990, Duncan and Humphreys, 1989). One fundamental point is that there are cases where searches for a single feature, like a unique colour (Nagy and Sanchez, 1990) or orientation (Foster and Ward, 1991b, Foster, 1991a), can produce steep set size slopes. This happens when the difference between the target and the distractor is decreased.

Another problem for FIT is the fact that some kind of conjunction search tasks produce flat slopes. Nakayama and Silverman (1986a, 1986b), found no effect of set-size on reaction time in searches for a conjunction of size and stereoscopic depth. Search asymmetries are another weak point for FIT. There is no set-size effect when searching for a tilted line among vertical lines, but when the target is vertical and the distractors are oblique, set-size effects can be quite significant (Treisman and Gormican, 1988).

Another weakness for FIT is the prediction of a 2:1 slope ratio for target present/absent searches. Such an assumption requires search to be complete, but, at the same time, that no item is checked twice. Moreover, it makes no prevision for errors. Chun and Wolfe (1996) suggested that many serial search are not completely exhaustive as predicted by the strict FIT model, but that they present error rates of about 5-10%. Most of these errors can actually be considered as "misses", suggesting that the search ended *before* all items were processed.

Finally, FIT assumes that only one item is attended at a time. This seems contrary to the idea of a ‘visual span’, that is the number of characters (or items) that can be recognized in a glance (Legge, 1997). Note that the visual span is thought to vary with viewing eccentricity (Pelli, 2007), (see also Chapter 2).

3.1.1.4 Beyond the dichotomy

The idea that searches can be divided into two classes, serial and parallel, is very attractive but it is not supported by the experimental data. Results of visual search experiments show more the evidence of a *continuum* of search results.

Cave and Wolfe (1990) proposed an alternative classification for visual search tasks to differentiate the easy ones, where the target "pops-out" of the display, from the *hard* ones, where search is slowed by an additional distractor. The former can be described as *efficient;* the latter *inefficient*.

3.1.1.5 The Guided Search Model

The lack of a clear dichotomy does not invalidate the idea that there are distinct mechanisms in visual search. For example, in Wolfe’s Guided Search model (GSM), the continuum can be explained by an early parallel mechanism working in tandem with a later serial mechanism (Wolfe et al., 1994).

Guided Search is probably the most influential modification of FIT. In the parallel stage, a topographic map is formed, where the activity elicited by each item depends on its similarity to the target and its difference from the distractors. In the next stage, *attention* is initially directed towards the location with the highest activation. Search terminates if a target is found, otherwise attention is directed to the next highest activation. Activation is a noisy process, so that the item eliciting the highest activation is not always the target. Observers are allowed to terminate the search before all items have been inspected. They have to *adjust* themselves depending on the task and on the set size.

3.1.2 Signal Detection Theory

3.1.2.1 Palmer's Max Rule

Independently of Cave and Wolfe, Palmer and colleagues (Palmer et al., 1993) also constructed a model of search based on a noisy parallel process. The difference was that Palmer et al. found no need for an additional, serial mechanism. Using Signal Detection Theory (Green and Swets, 1966), Palmer et al. were able to predict how set-size affected the threshold difference between target and distractors .

Consider, for instance, a simple feature task as searching for a tilted line amongst vertical distractors. Observers could simply check the response by using a noisy filter tuned to the target's orientation. Repeated presentation of this particular item will produce a bell-shaped distribution of responses. Because of the filter's tuning, vertical distractors would usually generate smaller responses than the target, but not always. The frequency with which a vertical distractor would generate a larger response than the target would depend on the target's tilt and on the amount of noise.

When target and distractor distributions are well separated (large difference in orientation), the target will always generate a larger response so that it is *easy* to detect the target. The task can be much harder when the distractors are not vertical, but slightly tilted in the same direction of the target. In this case, the distributions for target and for distractors would be overlapping. Palmer and colleagues' paradigm is characterised by brief displays and widely spaced stimuli (Palmer et al., 1993, Palmer, 1994, Palmer et al., 2000). Palmer further attempted to control for crowding (see Chapter 2) by using a spatial cue to define set size and preserving a constant number of items on the display. Performance thresholds were measured using a two-alternative forced-choice detection task.

As in GSM, Palmer's model assumes that decisions are based on the stimulus which elicits the maximum level of activation. The greater the set size, the higher the probability that the largest activation will be elicited by one of the distractors.

3.1.2.2 *Averaging*

Morgan, Ward and Castet (1998) tested the observers in a different task. The observer had to report the orientation (left or right of vertical) of the Gabor patch that appeared most non-vertical. For long exposure times (in certain cases until the subject responds), threshold tilt varied with the number of untilted distractors, just as predicted from Palmer's winner-takes-all spatial uncertainty model. However, for short exposures (100 msec) followed by a random dot mask, the rise in threshold with distractors was considerably greater than that predicted from spatial uncertainty. Adding a tilt to the distractors in the opposite direction to the target increased the thresholds further still. These results suggested the application of a different decision rule, in which the observer takes the mean of the noisy independent orientation values for all stimuli. This is the averaging model.

The same idea has been successfully applied to a similar task in the condition of crowding (Parkes et al., 2001). In Parkes et al.'s work, observers had to report the tilt of a central crowded Gabor presented in the periphery of the visual field. Note, unlike visual search, in crowding experiments, observers typically know which element is the target. In the authors' opinion, under crowded conditions, the orientation signals in the primary visual cortex are pooled before they reach consciousness. Consistent with this idea, the authors found that tilting the distractors in the same direction as the target decreased orientation thresholds, whilst tilting them in the opposite direction increased thresholds. Parkes et al.'s observers were unable to identify which of the Gabors were tilted, further supporting the idea that in crowding visual information is pooled together.

3.1.2.3 *Localisation*

Localisation would not be possible if observers only had access to a regional pool of orientation estimates. Solomon and Morgan (2001) studied localisation in uncrowded displays. They used several different exposure times and asked the observer to report both the orientation and the location of the target. Localisation was much better than chance, indicating that observers were not limited to a global pool of orientation estimates. However, for short exposures, items mis-identified as the target were more likely to be adjacent to the actual target than would be expected from a max-rule with independently encoded distractors. This proximity effect will be discussed further in the Appendix.

3.1.2.4 Speed-Accuracy-Trade-Off (SAT)

Many other authors have examined the dynamics of visual search. The most prominent of these are Dosher and colleagues, who have systematically examined what they call the Speed-Accuracy-Trade-Off (SAT). Like the paradigms discussed above, SAT is also amenable to analyses based on SDT. SAT provides conjoined measurements of discrimination (accuracy) and processing dynamics (speed). In particular this procedure shows how accuracy improves with processing time (Dosher, 1976, Dosher, 1979, Dosher, 1991). SDT was recently applied to SAT by (McElree, 1999). In McElree's experiments, processing time was controlled either by forcing observers to respond when a cue sound was presented or by limiting the exposure time.

The asymptote of the SAT function is a measure of target-distractor discriminability. The pre-asymptotic part depends on the processing speed. It shows how fast the performance grows and when the observer can do better than chance.

Consistent with Palmer and colleagues' findings, (e.g. 1993), McElree and Carrasco (1999) reported that an increase in the number of elements causes a decrease in target-distractor discriminability, however the speed of processing was only affected by set-size in conjunction searches; not in feature searches.

3.1.2.5 Salience maps in V1

Very recently, Li Zhaoping (2002) has taken Cave & Wolfe's (1990) idea of a topographic map, where activation depends on each item's similarity to the target and its dissimilarity from the distractors, and instantiated it with a biologically plausible model of V1. In this model items that produce particularly high activation are said to be "salient". Without evoking separate feature maps or any combination of them, this model is able to predict several psychophysical results. The model is particularly successfully in its account of search asymmetries (Li, 2002).

3.1.2.6 Conclusion

The safest generalisation about visual search is that the search time for a target decreases with its difference from the distractors, and with distractor homogeneity. We will see in the following Chapter that crowding can also be reduced by target salience. It is time to put these two sets of findings together, and to ask whether they are related. The next Chapter investigates the relationship between salience in visual search and crowding using the same stimuli in the two paradigms.

Chapter 4

The relationship between search efficiency and crowding

The role of target salience in crowding remains controversial largely because salience usually escapes objective measurement. Here we address this problem using search efficiency as a measure of target salience. In separate experiments, observers determined whether parafoveal arrays of vertical Gabor patterns contained targets having a unique colour, a unique direction of motion and a unique temporal frequency. We analysed search efficiency in the conventional manner using reaction-time gradients, with units ms/item. We also considered accuracy gradients, with units percent correct/item. Crowding is typically quantified by comparing the acuity for a target within an array to the acuity for a target presented alone. We measured orientation acuity for determining whether a slightly tilted target was clockwise or anticlockwise compared to the vertical. Targets with a unique colour or direction of motion were found to pop out (with one exception), i.e. reaction-time and accuracy gradients were insignificantly different from zero. Acuity for these targets was significantly greater than acuity for targets whose neighbours had the same colour and direction of motion. Manipulation of temporal frequency produced a wide range of search efficiencies. For three of four observers, we found a linear relationship between acuity and the accuracy gradient; shallow gradients being associated with high acuity. In general, we find that crowding is weakened when observers can find a parafoveally presented target quickly and accurately.

4.1 Introduction

As mentioned before, the deterioration of our ability to identify targets in the presence of distractors is referred to as *crowding* (Bouma, 1970, Stuart and Burian, 1962). Crowding tends to be greatest when target and distractors have similar stimulus properties, such as colour, shape and depth (Andriessen and Bouma, 1976, Kooi et al., 1994, Nazir, 1992). These studies appear to support Korte's opinion (1923) quoted by Pelli and Palomares (2004) that distractors 'squash' less salient targets. However, with one exception (discussed below), none of these previous studies actually quantified target salience.

In this study we compare crowding with an independent measure of salience: search efficiency (Wolfe, 1998). Salient targets are those that can be found quickly and accurately, regardless of how many distractors there are. We sought to determine whether the extent of crowding could be predicted from target salience. Salience was manipulated by changing the target's direction of motion, colour and temporal frequency. To determine the extent of crowding, we asked observers to discriminate between small changes in the target's tilt.

Previous research (Nakayama and Silverman, 1986a) has shown that a target moving in one direction will "pop-out" from distractors moving in the opposite direction; that is, such targets are demonstrably salient. Therefore, in order to be consistent with our hypothesis, such targets should be resistant to crowding.

Searches for a unique colour are also very efficient, given a sufficient difference between target and distractors (Nagy and Sanchez, 1990, Nagy et al., 1990, Green, 1956). Two of Kooi et al's (1994) six observers suffered crowding even when the target had a unique colour. Kooi et al. suggested that these two observers might not have experienced as strong a perceptual difference between the target and distractors. Thus we decided to re-examine whether a uniquely coloured target can be crowded, this time with an independent measure of target salience.

Finally, we sought to systematically manipulate target/distractor differences, thereby allowing for a whole range of search efficiencies. For this purpose, we selected a stimulus property, temporal frequency, that had not hitherto been investigated with respect to crowding.

4.2 Methods

Experiments were written in MATLAB, using the Psychophysics Toolbox (Brainard, 1997, Pelli, 1997). The stimuli were displayed with a Dell monitor (1024 x 768 pixels; display area 36.7 cm x 27.5 cm; frame rate 75 Hz; mean luminance 70 cd/m^2). Observations were carried out in a lighted room.

All observers had normal or corrected-to-normal visual acuity; three were the experienced observers (CG, MM, JAS), and the others were naive to the aim of the experiments (JK, MRC, SP, VA, GW, MOB). All data collected for each observer are shown below.

On each trial of each experiment, between one and nine stimulus items appeared in an array (see Fig. 4-1), randomly on the right or left of a point at the centre of the display. Observers were encouraged to maintain fixation on this point. At a viewing distance of 55 cm, the central item of each array appeared at 6.5 degrees of eccentricity. The stimulus configuration was a square. Outer items were centred either 1.09 or 1.53 degrees away from the centre of the central item.

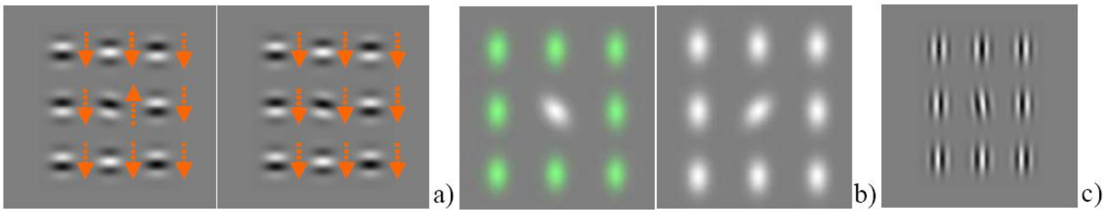


Fig. 4-1: Sample stimuli used to measure acuity in Experiments 1 (a), 2 (b) and 3 (c). Observers had to identify the direction of the central target's tilt, viewed at 6.5 degrees of eccentricity. In some conditions, the surrounding, vertical distractors were absent. When present, the distractors could have the same or different colour, direction of drift or temporal frequency as the target. When search efficiency was measured, all elements were vertical. Observers decided whether any one of them had a unique colour, direction of drift or temporal frequency. The orange arrows in fig 4-1(a) indicate the direction of motion of patches and did not appear in the experiment.

Acuity and search efficiency were measured in separate blocks. When the latter was measured, all items were oriented vertically. On half of the trials, an odd-man-out appeared at a random position. For example, this *target* item could have been green; the others (*distractors*) would have been white. The observer was required to decide "quickly and accurately", whether an odd-man-out was displayed. In Experiment 2, the display remained visible until the observer pressed a key to indicate his or her choice. In Experiments 1 and 3, the displays remained visible for fixed durations (2 s in Experiment 1; either 1.6 or 3.2 s in Experiment 3, see below for more details), which were longer than the longest reaction time recorded.

When acuity was measured, all items except the central *target* item were oriented vertically. On half of the trials, this target was also an 'odd-man-out' with respect to another dimension, such as colour. The observer was required to decide whether it was tilted clockwise (CW) or anti-clockwise (CCW) of the vertical. In Experiments 1 and 2, the display remained visible for 200 ms. In Experiment 3, the display was longer (see below). Observers were encouraged to be as accurate as possible, without time constraints.

When acuity was measured, the target appeared with one of six tilts (three CW, three ACW), preselected to produce well-sampled psychometric functions relating of the probability of a CW response to the size of the tilt, with ACW tilts counting as negative. The data were (maximum-likelihood) fit with two-parameter cumulative Gaussian functions using the MATLAB 'fminsearch' function. Analysing the data in this way made it possible to separate orientation acuity $(1/\sigma)$ from bias $(-\mu)$. Confidence intervals were estimated by a bootstrap procedure (Efron 1979, 1982). That is, maximum-likelihood estimates of μ and σ were used to define a simulated observer who performed the same experiment as our human observers 120 times. After each simulation, threshold and sensitivity were estimated anew, and the 95% confidence interval for threshold is bounded by the fourth largest and fourth smallest of these new threshold estimates.

When search efficiency was measured, accuracy and reaction time were estimated separately for target-present and target-absent trials at three set-sizes i.e. arrays containing three, five and nine stimulus items. Gradients were then estimated by regressing accuracy and reaction-time estimates against set size (e.g. Figs. 4-2 and 4-4).

4.3 Results

4.3.1 Experiment 1: Motion

In Experiment 1, each stimulus item was a maximum-contrast (88%) drifting Gabor patch with spatial wavelength $\lambda_x = 0.33^{\circ}$, temporal wavelength $\lambda_t = 0.067\ \mathrm{s}$ and space constants $\sigma_x = \sigma_y = 0.21^{\circ}$. Targets moving in a unique direction were found to pop-out. Specifically, neither reaction-time gradients nor percent-correct gradients were significantly different from zero for any observer ($p > 0.5$ except for JAS in the target-absent condition $p = 0.1478$; see Fig. 4-2). Accuracy exceeded 90% in all conditions.

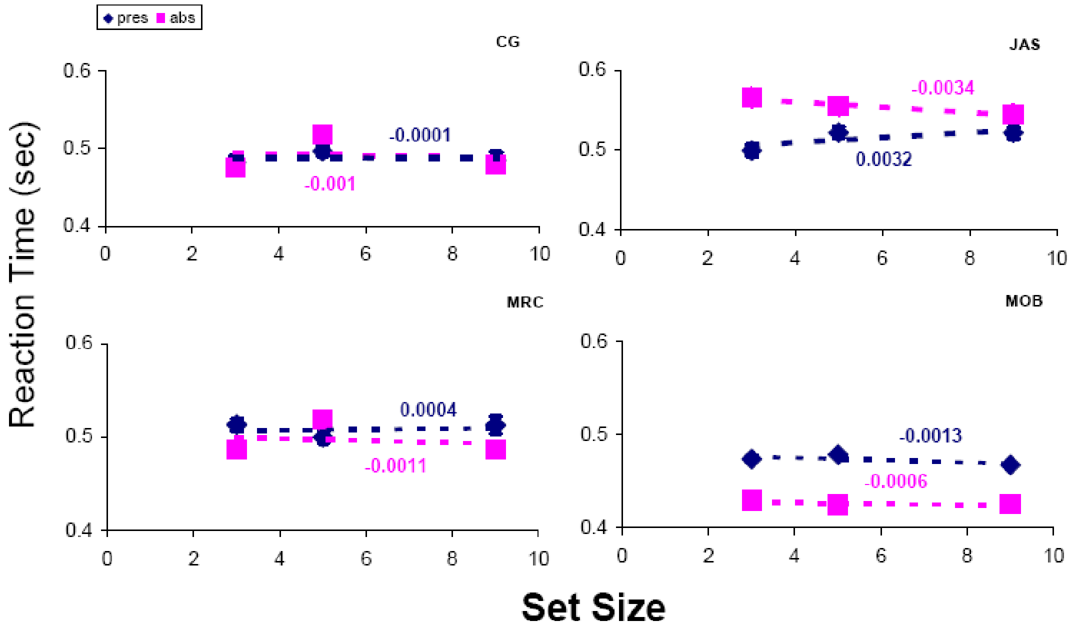


Fig. 4-2: Reaction times in searches for a unique direction of drift. Blue and pink symbols reflect target-present and target-absent conditions, respectively. Each data point summarises a minimum of 200 trials. Error bars contain the standard error. There was no significant increase in reaction time with set size for none of the four observers.

Acuities are illustrated in Fig. 4-3. Without distractors, baseline acuities ranged from 0.5 to 0.7 deg^{-1}. For each observer, acuity was significantly reduced ($p < 0.025$) when the target appeared with distractors moving in the same direction. When the target appeared with oppositely moving distractors, acuity improved significantly ($p < 0.05$), in some cases (JAS and MOB) reaching the baseline value.

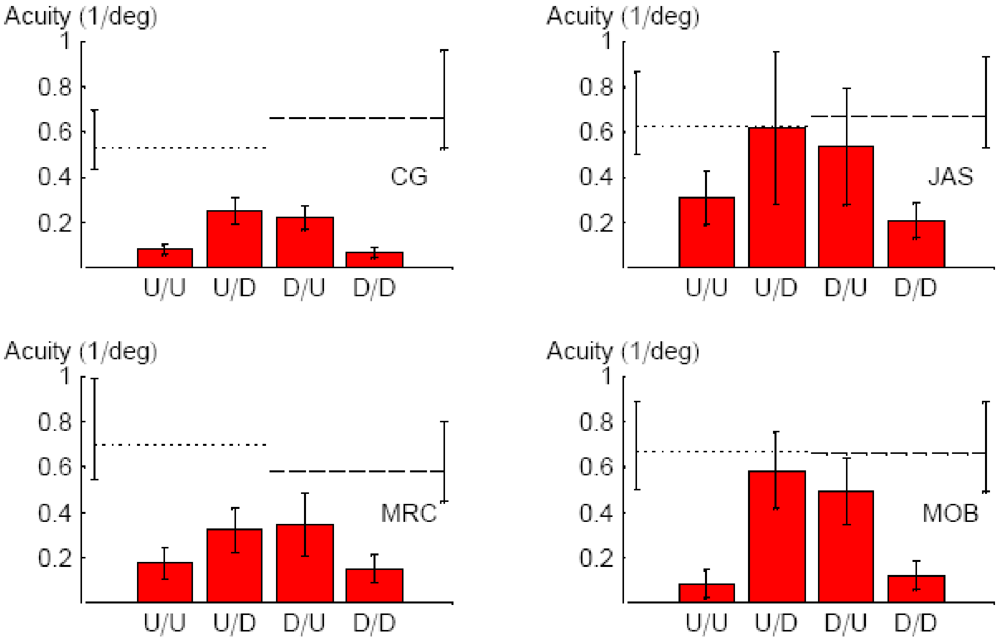


Fig. 4-3: Orientation acuities for CW/ACW judgments with respect to vertical for CG, JAS, MRS and MOB. Dotted and dashed lines reflect acuities for upward and downward drifting targets, respectively, without distractors. Different columns reflect different conditions, e.g. “U/D” is the condition in which the target drifted up and the distractors drifted down. Each column summarises a minimum of 120 trials. Error bars contain 95% confidence intervals. Acuity was worst when target and distractors shared the same direction of drift.

4.3.2 Experiment 2: Colour

In Experiment 2, each stimulus item was a white or green elongated Gaussian blob with large and small space constants $\sigma_y = 0.47^{\circ}$ and $\sigma_x = 0.31^{\circ}$, respectively. No attempt was made to look for an isoluminance point. The peak of the white blob had an RGB value of [255 255 255]. The peak of the green blob had an RGB value of [127 254 127]. The background had an RGB value of [127 127 127].

As in the previous experiment, uniquely coloured targets were found very efficiently. The largest reaction-time gradient was < 10 ms/item and only one gradient was significantly different from zero (CG, target present: p = 0.037; all others: p > 0.5; see Fig. 4-4). No percent-correct gradients were significantly different from zero (p > 0.5), and accuracy exceeded 90% in all conditions.

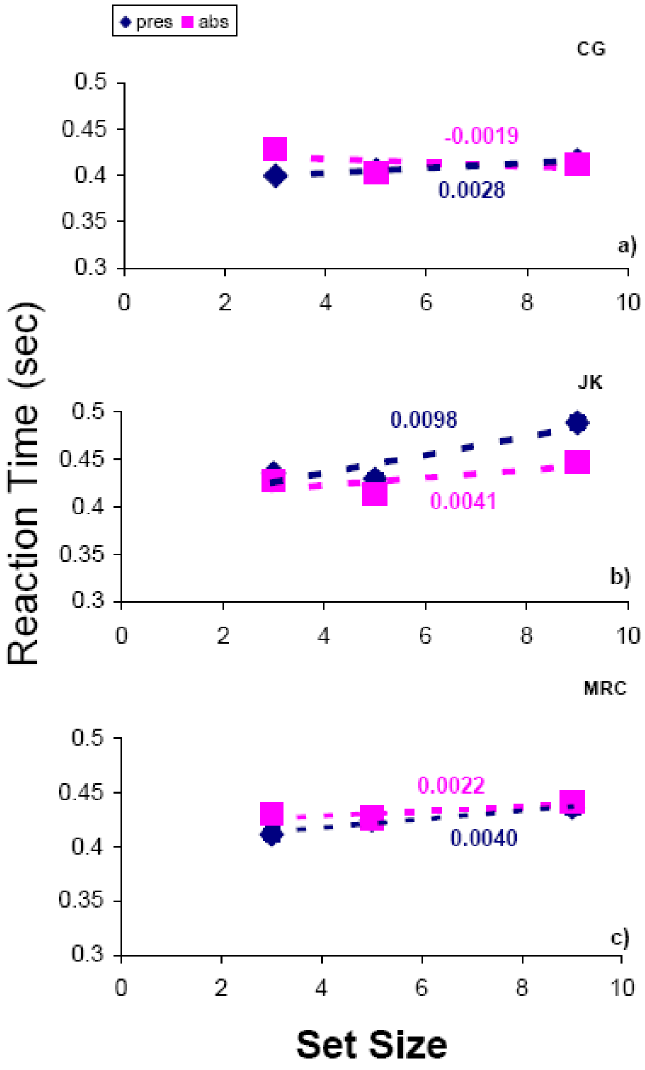


Fig. 4-4: Reaction times in searches for a unique colour. Blue and pink symbols reflect target-present and target-absent conditions, respectively. Each data point summarises a minimum of 220 trials. Error bars contain the standard error. Reaction time increased significantly with set size in one condition only (CG, target present; panel a).

Acuities are illustrated in Fig. 4-5. Without distractors, baseline acuities ranged from 0.2 to 0.3 deg-1. For each observer, acuity was significantly reduced ($p < 0.025$) when the target appeared with identically coloured distractors. When the target appeared with differently coloured distractors, acuity improved significantly ($p < 0.05$), in some cases reaching the baseline values.

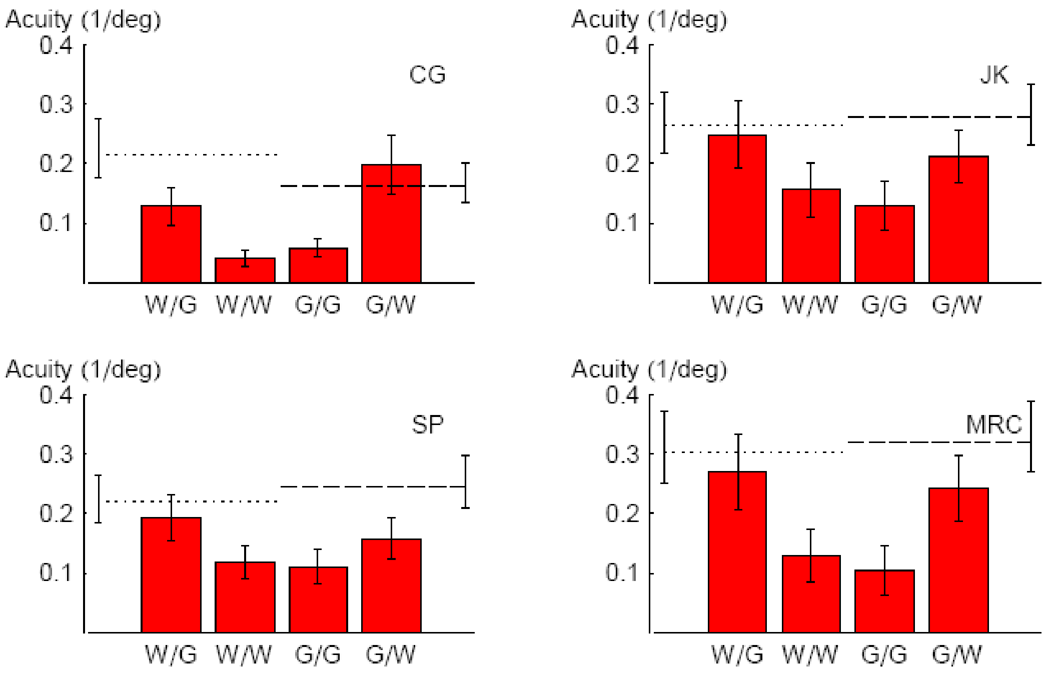


Fig. 4-5: Orientation acuities for CW/ACW judgements with respect to the vertical for four observers. Dotted and dashed lines reflect acuities for white and green targets, respectively, without distractors. Different columns reflect different conditions, e.g. “W/G” indicates white distractors and green target. Each column summarises a minimum of 360 trials. Error bars contain 95% confidence intervals. Acuity was worst when target and distractors shared the same colour.

4.3.3 Experiment 3: Temporal Frequency

In Experiment 3, each stimulus item was a (stationary, achromatic) Gabor patch ($\lambda_x = 0.53^\circ$, $\sigma_x = \sigma_y = 0.21^\circ$), whose contrast varied sinusoidally between 0 and *w*, at either 2.5, 7.5, 15 or 30 Hz. When acuity was measured, *w* was a Gaussian window, with time constant $\sigma_t = 200\ \mathrm{ms}$ and a peak contrast of 88%. The Gaussian was curtailed at $\mu \pm 2\sigma$, giving a total display duration of 800 ms. In this phase of the experiment, JAS performed a total of 2,742 trials. Each of the other observers performed more than 4,000 trials.

When search efficiency was measured, the temporal window *w* was a concatenated pair of the curtailed Gaussians described above. All observers except JAS were able to respond within the 1.6-s display. For JAS, four concatenated, curtailed Gaussians were used, producing a 3.2-s display. There were four target-absent conditions, one for each frequency; and twelve target-present conditions, one for each possible combination of target and distractor. All sixteen conditions were presented with equal frequency. Thus, when search efficiency was measured, the probability of a target in any given trial was 3/4. In this phase of the experiment, each observer performed a minimum of 576 trials.

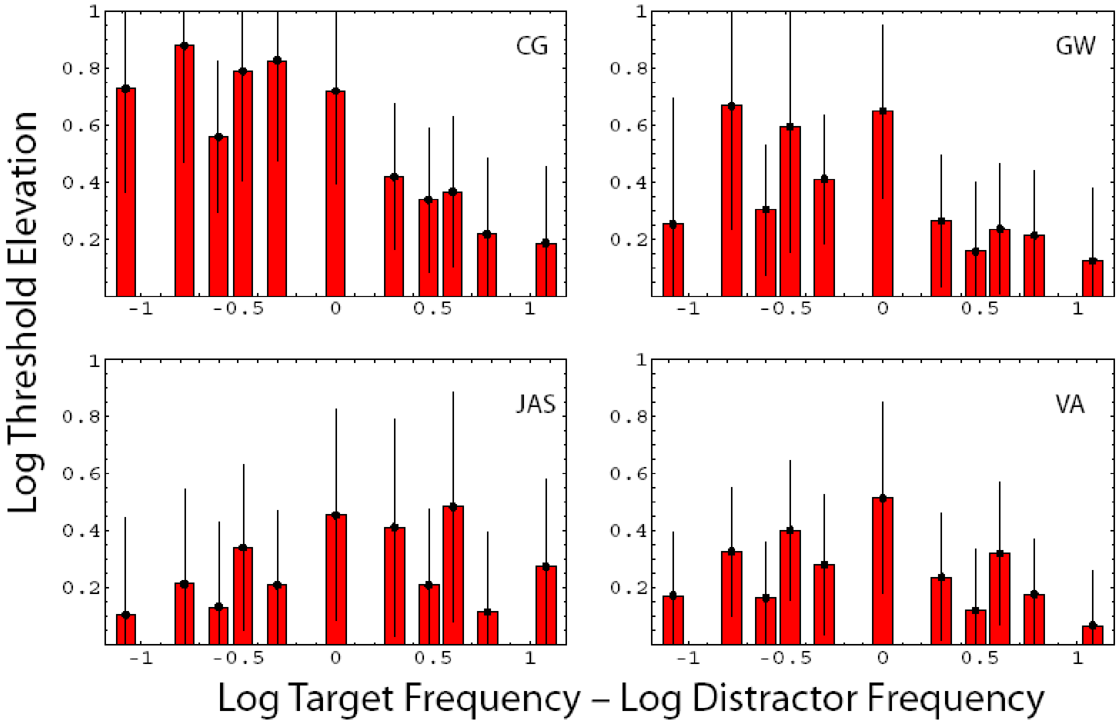


Fig. 4-6: Log threshold elevations for four observers. A value of 1 indicates that distractors produce, on average, a 10-fold drop in orientation acuity. Different columns reflect different ratios of target frequency to distractor frequency. Error bars contain 95% confidence intervals.

Without distractors, baseline acuities ranged from 0.4 to 0.7 deg^{-1}. For each observer, acuity was significantly reduced (i.e. thresholds were significantly elevated; $p < 0.05$) when the target appeared with distractors having the same frequency. We expected acuity to improve with the difference between target and distractor frequencies. As can be seen in Fig. 4-6, the results from observer VA are consistent with this prediction. However, for CG and GW, an improvement in acuity (i.e. threshold reduction) was consistent only when the target frequency exceeded the distractor frequency. For JAS, acuity improvement was consistent only when the distractor frequency exceeded the target frequency.

In Experiments 1 and 2, search accuracy exceeded 90% correct in all conditions. Experiment 3 was different in this respect. Some observers seem to have sacrificed accuracy for speed, failing to achieve even 50% correct in some conditions. For example, when the target frequency was 7.5 Hz and the distractor frequency was 15 Hz, CG's reaction-time gradient was a mere 1 ms/item; insignificantly greater than zero ($p > 0.6244$) and suggestive of a very efficient search. However, her accuracy significantly declined with set size ($p < 0.0367$), dropping approximately 4% with each additional distractor.

To get a better picture of the relationship between visual search and acuity, we simultaneously regressed log threshold elevations against reaction-time and percent-correct gradients. Specifically, we found the parameters b_0, b_r and b_p that minimised the sum of the squared errors $\sum_i e_i^2$ in:

$$\log E_i = b_0 + b_\mathrm{r} r_i + b_\mathrm{p} p_i + e_i ,$$

where r_i, p_i and E_i are the reaction-time gradient, percent-correct gradient and threshold elevation estimated from target-present condition *i*. The parameter b_r was not different from zero for any observer (CG: $p = 0.99$; GW: $p = 0.34$; VA: $p = 0.83$; JAS: $p = 0.53$). Therefore, we fixed the parameter $b_\mathrm{r} = 0$, and once again regressed threshold elevations against percent-correct gradients for each observer. The data are shown with the regression lines in Fig. 4-7. For three of four observers, the parameter b_p was significantly less than zero (CG: $p = 0.021$; GW: $p = 0.008$; JAS: $p = 0.049$), but for the fourth subject, b_p was not significantly different from zero ($p = 0.436$).

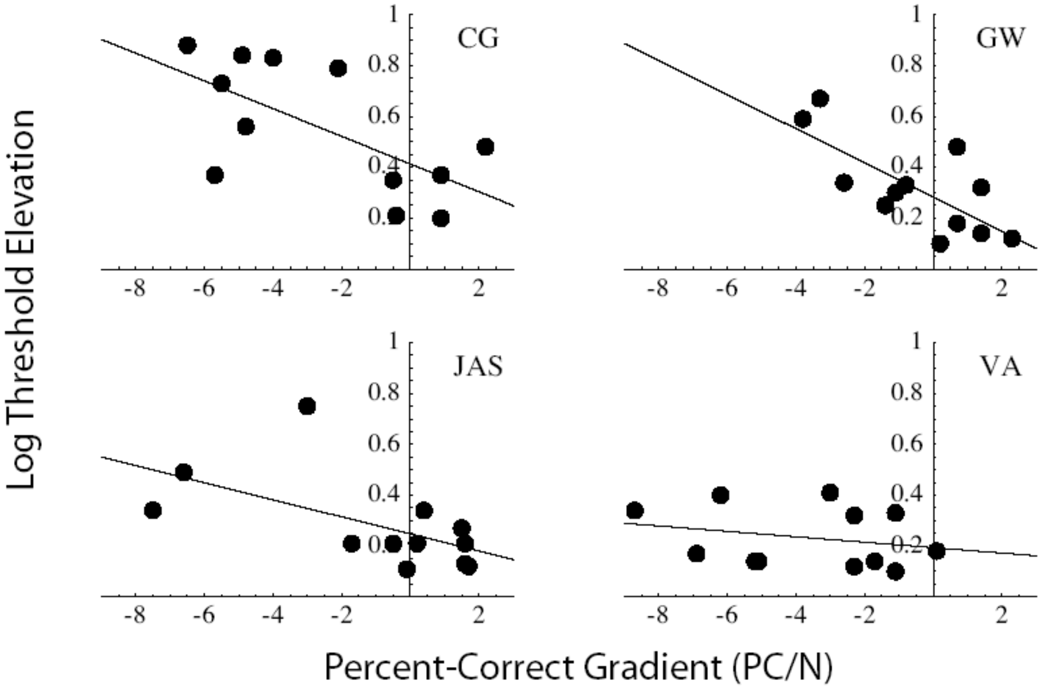


Fig. 4-7: The relationship between distractor-induced threshold elevation (i.e. crowding) and salience, for four observers. In each panel, each point represents one combination of target and distractor temporal frequencies. Conditions in which target and distractor frequencies were equal have be excluded from this figure. Log threshold elevation (i.e. the ratio between acuities for targets with and without distractors) has been regressed against the effects of set size (N) on reaction time percent correct (PC). Regression lines are shown. There was no significant relationship between threshold elevation and reaction-time gradient for any observer (not shown), but threshold elevation did increase significantly with the percent-correct gradient for observers CG ,GW and JAS.

4.4 Discussion

The results of Experiment 1 demonstrate that a target made highly salient because of its direction of drift is less susceptible to crowding than one in an array of similarly moving distractors. The results of Experiment 2 confirm a previous report (Kooi et al., 1994) that uniquely coloured targets can be made similarly resistant to crowding. Kooi et al also reported results from observers for whom target colour made no difference. Had we found such an observer, it would have been interesting to see whether his or her search efficiency was comparable to that of the other observers, but we did not. Experiment 3 did allow us to compare tilt acuities for targets of various saliences. Using a novel index of search efficiency (the percent-correct gradient), we found a strong relationship between target salience and susceptibility to crowding in three of our four observers. On the basis of these results, it seems safe to concur with Korte; more salient targets are resistant to being “squashed” by distractors. There is at least one exception to this rule, so glaring that it forced Felisberti et al (2005) to conclude that target salience had “at best, modest effects on crowding.” In a search task, Felisberti et al found that distractors having a 22.5-deg tilt with respect to the (vertical) target could more than double tilt thresholds for CW/ACW judgements with regard to the vertical (Solomon et al., 2004). *Prima facie*, these results seem to suggest that even very salient targets can be crowded.

However, further experiments have revealed a unique mechanism for grossly-oblique-distractor-induced acuity loss (cf. the threshold and sub-threshold distractor tilts used by Parkes et al. 2001), which may be impervious to manipulations of target salience. Because of the tilt illusion (Gibson and Radner, 1937), targets must be tilted toward oblique distractors in order to appear vertical. Gibson suggested that the labels “horizontal” and “vertical” shift from mechanisms preferring these orientations to those aligned with the visual context. Elsewhere (Solomon, 2006) we suggest that acuity suffers when the extent of this shift fluctuates from trial to trial.

We conclude that there may be several mechanisms responsible for crowding. One kind, seen with tilted distractors, depends upon the tilt illusion, and is found even when the target is highly salient. Another kind, seen with non-tilted distractors, depends on a different mechanism, such as obligatory pooling (Parkes et al 2001) and is reduced by increasing the salience of the target.

4.5 Conclusion

We have seen that colour is a dimension facilitating the visual search for an unique target. The purpose of the next Chapter is to investigate the claim that the same is true of the subjective colours experienced by grapheme-colour synaesthetes.

Chapter 5

Visual Search and Synaesthesia

Colour synaesthesia is an intriguing condition in which people have strong associations between visual or auditory stimuli such as numbers, and particular colours. For example, the number ‘2’ (spoken or written) may be experienced as ‘red’. These associations are idiosyncratic, but durable over time (Baron-Cohen, 1987). A key question for neuroscientists is whether these associations are purely verbal, such as the association that makes us think of ‘biscuit’ when we hear or read ‘Garibaldi’, or whether, more interestingly, they evoke experiences like those of real colours.

5.1 Introduction

The purpose of the present Chapter is to examine the suggestion by Ramachandran and Hubbard (2001b) that synaesthetic colours can aid visual search. In a previous Chapter (Chapter 3) we have shown that colour can help normal observers partially to overcome 'crowding', and we have linked this finding to the general concept of salience. Our original intention was to see if synaesthetic colours (that is, colours experienced by grapheme-colour synaesthetes when viewing objectively achromatic letters or numbers) could increase the salience of targets in visual search and crowding tasks, and if so, to probe further the underlying mechanism. We initially used a visual search task, because previous reports indicated that this is more robust than the crowding effect. However, to anticipate the results, we failed to find any effect of synaesthetic colours on visual search, so we did not go on the examine crowding.

5.2 Definition of the Condition

The word Synaesthesia has been used to describe an unusual state of perception. It derives from the Greek ***syn***, 'together' and ***aesthesis***, 'perception', that is 'joined sensation' and it is usually associated with a condition when a physical stimulation elicits a response in more than one sensorial modality. The synaesthetic experience is composed by two interrelated events: the *inducer* and the *concurrent* (Mark, 1975, Mills and el., 1999, Wollen and Ruggiero, 1983, Wollen, 1983). The first is the inducing event, the latter the synesthetically induced sensation. For example to a certain synaesthete who associates sound with colours, the barking of a dog (inducer) can be acid yellow (concurrent). For most individuals synaesthesia is *unidirectional*: following the same example, the colour yellow will not elicit the sound of the barking dog. Synaesthetes can describe 'sounds like coloured blobs' and they might state they can 'hear colours' , 'taste voices' or 'smell colours'. Popular literature is replete with myths about the condition. The typical synaesthete has been described as a rare left-hand woman (Cytowic, 1989, Baron-Cohen, 1993), artistically creative, but with poor mathematical and topographical ability and as artistically gifted (Galton, 1880, Seron et al., 1992, Marks, 2000, Baron-Cohen et al., 1996).

Modern research has questioned these myths. Synaesthesia may be more common than has been supposed but recent studies have confirmed that it is more frequent amongst the female population (Rich and Mattingley, 2002). Rigorous measures of synaesthetes' personality do not exist to support claims that there is a specific synaesthetic personality. Some synaesthetes may have an abnormally good memory, such as Luria's case (Luria, 1968), but there is no evidence that this generalises to the population of synaesthetes.

It is generally agreed that Synaesthesia is automatic and involuntary (Cytowic, 1996). The experience is accompanied by a sense of certitude and a conviction that what synaesthetes perceive is real and valid. They talk about their perception as if it were obvious and 'normal'. They are actually surprised when they find out for the first time that what they perceive is not shared by the majority of the population .

Synaesthesic concurrents can be either projected (Dixon, 2000, Cytowic, 1996) or non- projected(Grossenbacher and Lovelace, 2001). If it is projected, the concurrent is perceived externally in peri-personal space, the space immediately surrounding the body. For example, some synaesthetes report that they perceived the year as a coloured ring all around their body, each part representing a different colour and a different month. If it is not projected it is internally represented. For example, some colour-grapheme synaesthetes when presented with numbers, report seeing the colour 'in their mind's eye', rather than in external space (Cytowic, 1989, Grossenbacher and Lovelace, 2001, Cytowic, 1993).

An important property of synaesthesia is that the association between inducers and concurrents are idiosyncratic, but durable over time (Baron-Cohen, 1987). Baron-Cohen et al (1987) asked synaesthetes to report their synaesthetic colours for 103 words. Words include animals, places names, objects, abstract terms, days of the week, people names and letters. Observers were retested, without warning, 10 weeks later, synaesthetes were 100% consistent in the colours they reported for the entire set. By contrast, the non-synaesthetic controls were able to recall 17% of the colour associations they provided for the 103 items after just two weeks, even if they had been encouraged to use memory strategies for associating words with colours and to remember them. Such tests of consistency have become the 'gold standard' for determining the 'genuineness' of synaesthesia (Cytowic, 1997, Baron-Cohen, 1996, Baron-Cohen, 1993, Cytowic, 1989, Mattingley et al., 2001, Rich and Mattingley, 2002, Ward, 1993), although it has been noted that in some cases synaesthetic experiences can change over time (e.g. become weaker with age).

A very central question is whether synaesthetes are a homogeneous group. Some investigators have suggested two sub-groups: ‘higher synaesthetes’, in whom the cross activation occurs between ‘higher colour areas’; and ‘lower synaesthetes’ where quite early visual areas are activated (Ramachandran, 2003, Hubbard, 2005b). It has also been suggested that only some kinds of synaesthetes, called ‘projectors’ actually see their synaesthetic colours projected into the outside world (Smilek, 2001). More recently, Ward et al (2007) compared the behavioural performance of 7 projector and 7 associator synaesthetes, and showed that this distinction does not map on to the higher–lower distinction. They replicated previous research showing that projectors are faster at naming their synaesthetic colours than veridical colours, and that associators show the reverse profile. Synaesthetes who project colours into external space but not on to the surface of the grapheme behave like associators on this task. In a second task, graphemes presented briefly in the periphery were more likely to elicit reports of colour in projectors than associators, but the colours tended to be accurate only when the grapheme itself is also accurately identified. Ward et al. (2007) proposed an alternative model of individual differences in grapheme-colour synaesthesia that emphasises the role of different spatial reference frames in synaesthetic perception.

One of the major issues in synaesthesia research is, in fact, to identify the level of processing involved in the formation of the subjective colours experienced by synaesthetes: are they low-level sensory phenomena, equivalent to real colour signals, or are they due to memory and association learning? To address this question, we tested whether the colours reported by a group of grapheme-colour synaesthetes (previously studied in an fMRI experiment) influenced them in a visual search task. Unlike previous studies of the same problem, we also introduced a condition where the colours experienced by synaesthetes would be expected to make them worse than controls. We found no evidence for differences between synaesthetes and normal controls, either when colours should have helped them or where they should have hindered. We conclude that the colours reported by our population of synaesthetes are not equivalent to low-level sensory signals, but arise at a cognitive level where they are unable to affect visual search.

5.3 Review of the evidence

A recent paper on synaesthesia (Kim, 2006) states that ‘based on recent work in a number of laboratories it is *now impossible to dispute* that at least some forms of synaesthesia entail mental experiences that have a genuine perceptual reality’. The work to which this statement refers claims to show that the perceptual colours experienced by synaesthetes have equivalent effects to real colours in psychophysical tasks. To say that it is ‘impossible to dispute’ this conclusion is a very strong claim, which nevertheless, we shall challenge in this chapter. We shall first analyse the existing evidence, and then present an experimental investigation designed to defeat the expectation that synaesthetes should show superior performance to normal controls.

A brief introduction may be required to explain the logical issues involved in testing unusual subjects like synaesthetes. It is usual in psychophysics to distinguish between purely phenomenological reports and measures of performance. The reports by synaesthetes seeing certain digits as coloured is purely phenomenological. So is the observation by normal observers that they see a previously ‘yellow’ patch as tinged with red after adapting to a green patch. Phenomenological measures have a long history and an important role to play in the study of perception. However, a potential problem with them is that they are easily influenced by expectations and biases. If we encountered an unusual person who persistently called ‘red’ stimuli ‘green’, we might wonder whether there was something genuinely unusual about their experience, or whether they were using words in an idiosyncratic way, possibly to attract attention to themselves. We would therefore seek a *performance* measure to supplement the verbal report. For example a genuinely colour deficient observer would be expected to be poor in identifying the numerals in certain plates of the Ishihara test. If they performed poorly with the plates where we would expect poor performance due to the absence of abnormality of one class of cone receptor, we would be assured that the deficiency was real. If the observer failed on all the plates we would suspect that they were an impostor. Unfortunately, it is possible to feign a poor performance on certain plates if you are well-informed, to give the impression of colour vision deficiency. Performance tests are not infallible, except in the case where the unusual subject performs better than the normal population: superior performance is impossible to feign. If this is not the case, we have to be a bit cleverer. One of the Ishihara Plates makes use of the phenomenon of ‘colour camouflage’ to obtain a situation where the colour-deficient observer should have superior performance to the normal observer. It is much harder for an observer to feign a condition if they do not know whether they are supposed to be better or worse than the normal observer. This is the logic we shall use in the experiment described below.

The evidence that synaesthetic colours have similar perceptual properties to normal colours can now be assessed. Briefly, we claim that there is no satisfactory evidence that synaesthetic colours produce an objectively measured increase in performance relative to normal controls. In these circumstances we cannot exclude the possibility that the results given by synaesthetes reflect a natural, probably an unconscious, bias towards behaving as if their synaesthetic colours were real, supported by experimenter expectations (since no blind testing has yet been done, to our knowledge). Crucially, in this literature, far greater effort has been devoted to the collection of data tending to support the reality of synaesthetic colours than to trying to disprove it. In particular, the logical equivalent of the Ishihara test has not been reported until now. The one attempt to turn the synaesthetes’ supposed ability against them (the Stroop effect, see below) only succeeds in showing semantic, not perceptual, interference, as even supporters of the reality of synaesthetic colours allow. The following review of the evidence is not comprehensive, but is intended to illustrate the main kinds of experimental evidence, and their strengths and weaknesses in addressing the question of the reality of synaesthetic colours.

5.3.1 The Stroop Effect

Several studies have shown that synaesthetes are impaired by ‘Stroop Interference’ (Mills and el., 1999, Odgaard, 1999, Wollen and Ruggiero, 1983, Dixon, 2000).

In the standard Stroop paradigm colour names such as RED are printed either in a congruent ink colour (red) or an incongruent colour (green) and subjects are asked to name the colour of the ink. Responses in the incongruent colour are slower than in the congruent. In the ‘synaesthetic Stroop’ paradigm, graphemes are presented in colours that are either congruent or incongruent for each synaesthete. Colour naming times are slower in the incongruent condition. Because this experiment uses synaesthetes as their own controls, it is not subject to the criticism that they may be more motivated than control subjects, or that they wish to demonstrate that they have special expertise. However, the Stroop experiment, does not prove that the colour associations are perceptual. The interference could occur at a semantic level, where two response tendencies, one evoked by the real colour, and the other by the association with the grapheme, collide (Hubbard, 2005a, MacLeod, 1988, MacLeod, 1991). As Palmeri (2002) concluded: ‘Indeed, such Stroop interference could occur even if synaesthetic colours have no perceptual reality whatsoever.’

5.3.2 The Incongruent Background Effect

Smilek (2001) asked a single synaesthetic subject (known by initial ‘C’) and normal subjects to name a digit presented on a coloured background. If the digit had the same synaesthetic colour as the background (‘congruent’), the synaesthetes were significantly slower to name it correctly than if it had a different colour (‘incongruent’). This is a curious result in several ways. First, it is unlikely that the background was exactly the same as the synaesthetic colour, in which case there would have been some residual colour contrast even in the ‘congruent’ condition. No evidence was presented that *normal* subjects were slower at identifying digits at low rather than high chromatic contrast. Second, the result implies that the activation threshold for verbally reporting a digit is higher than for evoking its synaesthetic colour In other words, the observer has to identify the digit as ‘4’ in order to see it as ‘red’, and then seeing it as ‘red’ slows down the ability to name it as a ‘4’ against a red background. Smilek et al. suggest a speculative feedback model to explain this paradox, but no demonstration that the model has been instantiated and that it actually works. Finally, and crucially, no evidence is presented that C did significantly better in the ‘non-congruent’ condition than the normal observers; indeed, she did worse than at least one normal observer. This means that there is no evidence for a sensitivity difference, and that C could have produced a congruent vs incongruent difference by unconsciously slowing her response to the congruent stimuli (a bias). In a second task, C localised one of two possible digits (2 or 4) when presented among 6, 12 or 18 distracter digits (a group of 8's). As in the first task C was faster on incongruent than on congruent trials. Once again, however, C was no faster on incongruent trials than at least some of the normal controls, leaving open the possibility that the effect is due to a slowing of response on congruent trials. It may also be doubted whether the difference between C and the normal controls reached conventional levels of significance. The difference was significant for C and not for any of the 7 normal controls, but it would be necessary to show that this difference is itself significant, by a rank order or similar nonparametric test, and no such test is reported. In a later experiment, Smilek et al (Smilek, 2001) an alphanumeric-colour synaesthete, J, and seven non-synaesthetes searched for target digits presented against backgrounds that were either congruent or incongruent with the colours of J's photisms for the target digits. For J, the slope of the search function for detecting the target digits on incongruent trials was shallower than the slope of the search function for detecting the target digits on congruent trials. In contrast, for the seven non-synaesthetes, the slopes of the search functions for detecting the target digits on congruent and incongruent trials were equivalent. Again, this result could have arisen from a proportional slowing of response on congruent trials.

Since the performance measure in the incongruent background effect is the naming of a digit, there is also a possibility that the locus of interference is linguistic, as in the Stroop task. Shown a digit, a synaesthete may have two competing tendencies, for example, ‘two’ and ‘green’. If the background also evokes the name ‘green’, it might reinforce the colour-naming tendency over the digit-naming tendency. This explanation may seem unlikely, but it is no less likely *a priori* than the idea that a congruent background reduces colour contrast of the digit.

5.3.3 Visual Search

If normal subjects try to find a ‘2’ amongst a set of ‘distracter’ digits such as 5’s their performance deteriorates as the number of distracters increases (the set-size effect). However, if the target is uniquely coloured it ‘pops out’ and performance shows little change with distracter number(Treisman and Gelade, 1980). Visual search with both real and synaesthetic colours was investigated by Palmeri (2002) in a synaesthete WO, along with normal controls. WO responded more quickly than the mean of the controls, when the search item had a unique synaesthetic colour. WO was not significantly better than controls when target and distracters were a similar synaesthetic colour (bluish), which is good evidence against a motivational explanation, provided that the different shapes were randomly interleaved rather than blocked. WO showed a clear set-size effect even with synaesthetic colours, leading Palmeri et al. to conclude that “the pop-out effect reported by WO may not be exactly analogous to the pop-out experienced with real colour”. In contrast, (Laeng, 2004) reported a single-case study where the synaesthete PM showed virtually no set-size effect for 2 amongst 5’s when the target had a unique synaesthetic colour. However, further analysis showed that fast response times occurred only when the target was within a few degrees of visual fixation, and the authors conclude that “PM’s synaesthesia does not occur preattentively, but rather is within the focus of attention.”

Edquist, Rich, Brinkman, & Mattingley (2006) presented data from 14 grapheme-colour synaesthetes and 14 matched non-synaesthetic controls, each of whom performed a visual search task in which a target digit was distinguished from surrounding distracters either by its unique synaesthetic colour or by its unique display colour. Participants searched displays of 8, 16 or 24 items for a specific target. In the chromatic condition, target and distracter digits were presented in different colours (e.g., a yellow '2' amongst blue '5's). In the achromatic condition, all digits in the display were black, but targets elicited a different synaesthetic colour from that induced by the distracters. Both synaesthetes and controls showed the expected efficient (pop-out) search slopes when the target was defined by a unique display colour. In contrast, search slopes for both groups were equally inefficient when the target and distracters were achromatic, despite eliciting distinct colours for the synaesthetes under normal viewing conditions. Edquist et al. concluded that, at least for the majority of individuals, synaesthetic colours do not arise early enough in visual processing to guide or attract focal attention.

Sagiv, Heer, and Robertson (2006) tested the reality of synaesthetic colours in two synaesthetes who perceived greyscale letters and digits in colour. They found no evidence for preattentive binding using a visual search paradigm in which the target was a synaesthetic inducer. In another experiment involving colour judgments, they showed that the congruency of target colour and the synaesthetic colour of irrelevant digits modulates performance more when the digits are included within the attended region of space. Sagiv et al. (2006) propose that the mechanisms giving rise to this type of synaesthesia appear to follow at least some principles of normal binding, and that synaesthetic binding seems to require attention. Their suggestion agrees with the finding (Laeng, 2004) that synaesthetic facilitation of search occurs only when the target is within a few degrees of fixation.

Studies of visual search, thus, have failed in general to find evidence for a pre-attentive effect of synaesthetic colours. There is no convincing evidence from this source that synaesthetic colours are perceptually equivalent to real colours. No studies so far have attempted to slow up visual search by synaesthetic colours using the principle of colour camouflage (Morgan et al., 1992) .

5.3.4 Visual Grouping

A set of black 2's forming a global shape such as a triangle does not, for normal subjects, stand out from a background of 5's (Beck, 1982). However, if the 2's are red and the 5's are green, the global shape is seen much more easily. Ramachandran and Hubbard (2001b) investigated whether this was true for synaesthetic colours as well. They reported that two synaesthetes were significantly faster than 40 control subjects at reporting the shape formed by achromatic graphemes that had a different synaesthetic colour from the background graphemes. The experiment was later extended to six synaesthetes (Hubbard, 2005b), five of whom showed superior performance to their control groups (n=20). An overall Analysis of Variance (ANOVA) yielded a significant effect of group (synaesthetes vs. controls). However, it is interesting to note that one of the control groups was clearly significantly superior to the other five, suggesting some caution in generalising to the population. If an arbitrarily chosen sample of 20 subjects can be significantly different from 100 other subjects in this task, then the chances of getting 6 synaesthetes that are different purely by chance must also be considered.

These two studies appear to offer convincing evidence that synaesthetes can do better than controls in a visual grouping task. But does this necessarily mean that they used their synaesthetic colours, rather than that they were trying harder and were more motivated? Clearly, the use of appropriately matched controls is very important in this context. In the experiment by Hubbard (2005b) the controls were undergraduate students carrying out the experiment for a course credit. It may be questioned whether they were giving the task as much attention as a group of synaesthetes keen to demonstrate their special gift. Another problem is the possibility of perceptual learning. Although the discrimination of graphemes in textures is difficult it may well improve with practice, as do other kinds of texture segregation (Casco, 2004). The synaesthetes, if they were giving the task greater attention than the controls, might have been able to speed up their performance by learning. It would be interesting to contrast their performance on early and later blocks of the 8 block (x 32 trials) experiment.

A final question is whether the superior performance of the synaesthetes, if it is not due to a motivational factor, necessarily implicates a pre-attentional effect of synaesthetic colour. The display was brief (1 sec) but possibly long enough to find several graphemes comprising the shape in a serial search. If this is so, the colours helping the synaesthetes to bind the graphemes into a shape may have been post-attentional. This possibility could have been tested by varying the number of distracters in the task, however this was not done. As in the case of visual search, therefore, it is possible that the shape did not truly ‘pop out’ for the synaesthetes.

5.3.5 Apparent Motion

Kim et al. (2006) reported results of an investigation in which real colours biased the direction in which observers saw apparent motion in a 2-frame movie sequence. In two synaesthetic subjects (WO and LR) their synaesthetic colours also biased perceived motion direction. Kim et al. (2006) conclude that synaesthetic colours are present at a stage of processing before the matching of tokens for movement perception, but this conclusion is not forced by the data. In fact, since what was measured was a bias, the effect could have arisen at any stage up to the final decision of which button to press. In a two-button forced-choice task, a subject has only to adopt the rule 'if in doubt press the left-hand button' to produce a shift in the psychometric function that will emerge as a bias. The fact that there may have been some complicated strategies was indicated by WO's performance when real colours were present in the task. In contrast to normal subjects, WO showed no effect of colour on matching, a fact which the authors were unable to explain. It would be informative to carry out an experiment on a normal subject who was rewarded by a points system for replicating WO's bias. Such experiments are standard in signal detection theory (Green and Swets, 1966).

5.3.6 Crowding

Ramachandran and Hubbard (Ramachandran, 2001a, Ramachandran and Hubbard, 2001b) reported that synaesthetic colours could also reduce 'crowding'. Graphemes presented in the periphery of vision are harder to identify when they are surrounded by other, flanking graphemes. Two synaesthetes were better than controls at identifying the target grapheme when it had a different synaesthetic colour. Subsequent research with six synaesthetes found the superior-to-normal effect in three of the six (Hubbard, 2005b). This is an extraordinary result, since it seems to imply that graphemes have already been identified before the site of crowding. If this is so, why is there 'crowding' for achromatic graphemes? The same objection does not apply to studies where crowding is relieved (to a small extent, it should be said) by real colours (Kooi, 1994, Gheri, 2008) because in that case, colour could increase target salience without prior target recognition. The crowding result is so counterintuitive and hard to explain on rational grounds, that it could serve as a useful focus for replication. If it cannot be replicated, the case would be strengthened that the synaesthetes in the Ramachandran and Hubbard study were somehow responding to their own and the experimenters' expectations.

5.4 Conclusions from the evidence

The evidence for a low-level colour input from graphemes in synaesthesia is, we propose, suggestive but far from convincing. The field has suffered from lack of a rigorous effort to rule out experimenter expectations and observer bias as explanations for the data. Apart from the Stroop test, which does not address to the issue of the perceptual reality of synaesthetic colours, there has been no determined effort to design a task where synaesthetes should be at a disadvantage relative to normal controls. In an attempt to do this we designed a task in which the presence of colour has been shown to interfere with visual search (Callaghan, 1984, Morgan et al., 1992) . In this task, search for a unique shape is impeded by randomly colouring the texture elements. This is not because of luminance differences between colours since the colour variation has no detrimental effect for dichromats (Morgan et al., 1992); rather it seems that colour is such a strong cue for textural segmentation that it impedes organisation based on other attributes, such as shape. We therefore expected to find that synaesthetes would be impaired in a search task by both real and by synaesthetic colour. Normal subjects, not having the apparent colours, should be in the same position that dichromats were in the Morgan et al. study (1992).

The synaesthetes were a sub-group of the subjects previously tested in an fMRI study by Nunn et al. (2002) and characterised as having grapheme-colour synaesthesia on the basis of a standard test (Baron-Cohen, 1996). We tested controls vs. synaesthetes with achromatic numerals that had previously been shown to have consistent subjective colours for the synaesthetes, to determine whether their subjective colours would interfere with visual search.

5.5 Methods

5.5.1 Measuring and verifying the subjective colours of synaesthetes

During the course of a previous investigation (Nunn et al., 2002) our synaesthetic subjects used swatches of paint to reproduce as exactly as possible the hues of each of the numbers 0-9. The key to our methodology in Experiment 2 is that some numbers had the same perceived hue associated with them, while others were different. It was thus important to determine that the hue associations were stable over time. We therefore scanned the paint swatches previously produced by the synaesthetes and used Adobe Photoshop to determine their RGB values on a VGA flat screen.

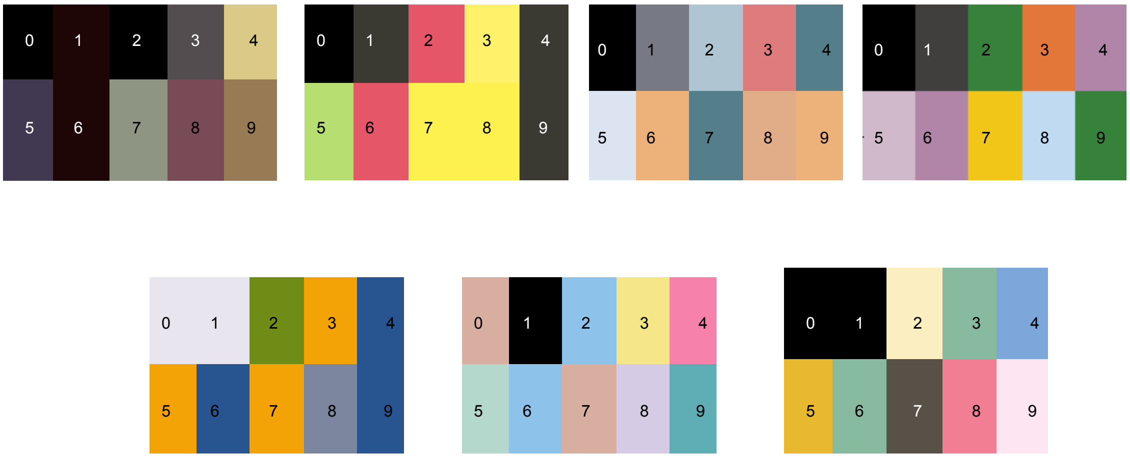


Fig. 5-1: Each subject matched their perceived colour of the digits 0-9 to a painted swatch and the latter were converted to RGB values for the experiment.

We then presented each subject with a matrix of swatches having the full gamut of RGB values associated with the numbers 0-9 and asked them anew to indicate the number that they associated with each colour (see Fig. 5-1). Their choices reproduced their previous associations without error, and furthermore, all subjects agreed that the colours on the screen were a good match to their colour associations.

It should be noted at this point that we never presented these colours during either experiment, and we do not claim that these colours were exact matches to the subjective colours. In Experiment 2, the crucial condition, all the numbers were achromatic (black). The purpose of verifying the colour associations was to show that they were stable, and to be able to present number pairs (for example, 2 and 8) which had the same subjective colour for a particular subject.

5.5.2 Experiment 1

The purpose of the first experiment was to see whether ‘real’ colours could help visual search for an unique object, both in synaesthetes and normals, and to replicate the colour camouflage effect reported by Callaghan (1984, Callaghan, 1986) and by Morgan et al (1992). Before starting the experimental session each subject underwent a training session until they felt confident with the task. Participants were not informed about the reasoning behind the experiment, and it is unlikely that any synaesthete could have inferred that they were expected to do worse than controls in the crucial test condition.

There were three experimental conditions (see Fig. 5-2). The first two were randomly interleaved within a block of trials. The stimulus was a 4×4 matrix of different coloured (red, green, blue, purple) and oriented lines, 16 in total, presented in the centre of the screen (each line 0.8°× 0.2°; total stimulus size 4.5°× 4.8°). On each trial a line was unique in colour and another was unique in orientation. A written cue was shown for 500 ms before the matrix randomly indicating whether the target was the line with the unique orientation or the one with the unique colour. The subject had to indicate whether the target was in the top or the bottom half of the display by pressing a button on the keyboard. In the third condition the lines were black and the subject had to look for the line with unique orientation.

Every subject collected 60 trials for each condition. We measured reaction time and percent correct.

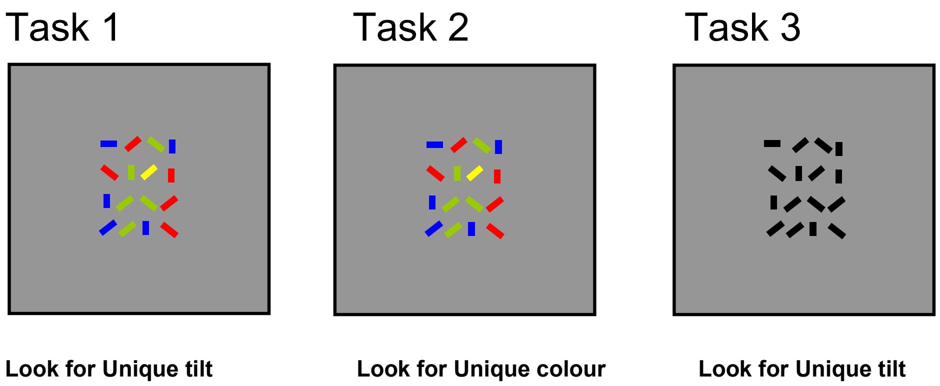


Fig. 5-2: Example of the screen for the Experiment 1. There were three experimental conditions. The task of the observer was always to indicate the location of the target. In the first two a written cue ‘shape’ or ‘colour’ indicated if the target was the line with the unique tilt or with the unique colour (0.8°× 0.2°; total stimulus size 4.5°× 4.8°). A third condition was programmed were lines were black. The observer in this case had to indicate the line with the unique tilt.

5.5.3 Experiment 2

In this Experiment we used numbers instead of lines. The stimulus was a 4×4 matrix of different numerals (Fig. 5-3) presented in the centre on the screen. The numbers presented were all black and they were all repeated at least once except for the target. (For example, in the set 3 6 6 7 6 7 5 6 3 7 5 6 3 5 7 8, the number 8 is the target).

Two conditions were prepared for synaesthetes. In the first (the ‘unique condition’) the numbers for each synaesthetic subject were chosen such that the target, as well as being an unique number, also had a unique subjective colour for that subject. In other words, all the distracters had the same subjective colour for that subject, which was different from the colour of the target. In the second condition (‘non-unique’) the target shared its subjective colour with that of at least one of the distracters. Every synaesthete was paired with an age-matched control who was presented with exactly the same stimuli and in the same order. The unique and non-unique conditions were randomly interleaved and 60 trials were collected in each condition. Constraints on the availability of differently-coloured numbers for the synaesthetes meant that the unique condition could contain only 5 different numbers, while the non-unique contained 6 different numbers.

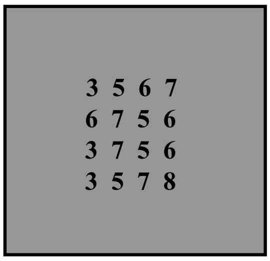
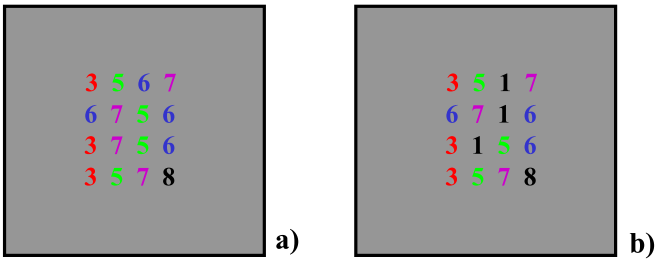


Fig. 5-3: Example of the stimulus for experiment 2. In this test all the numbers of the matrix were black (0.8°× 0.5°; total stimulus size 4.5°× 4.8°; spacing between numbers 1°). The task was to indicate the location of the unique number. Two conditions were programmed depending on which colour subjects would associate to each number: the unique one, where the target was the only item with a certain colour and the no-unique condition where the perceived target colour was repeated on different numbers.



Unique colour condition **No Unique Colour condition**

Fig. 5-4: Example of the screen for experiment 2 as it may look to a specific synaesthete in the two different conditions: Unique colour condition a) and No Unique colour condition b).

5.6 Results

To obtain a normal distribution of data and to be able to use parametric statistics all the reaction time data were logarithmically transformed.

5.6.1 Experiment 1

Percent correct was high for all subjects in all three conditions and showed no significant difference between conditions. Mean reaction time results for experiment 1 are shown in Fig. 5-5. Both synaesthetes and controls show the same effect: looking for the unique coloured line was easier than looking for unique tilt. More interestingly looking for a unique tilt was faster when all the lines were all black than when they were differently coloured (Condition 3 vs. Condition 1), confirming the previous result by Morgan et al. (1992).

An ANOVA test found no significant difference between the two groups ($F = 0.636$, $p = 0.430$) and no significant interactions ($F = 1.744$, $p = 0.189$). A significant difference was found between conditions within both groups (synaesthetes: $F = 74.427$, $p < 0.01$; controls: $F = 86.39$, $p < 0.1$). A series of Tukey post hoc tests showed that in both groups all three conditions were significantly different from each other (Synaesthetes: $p < 0.01$; controls: $p < 0.01$).

A possible explanation of the difference between conditions 1 and 3 is that the former was blocked with the colour-search condition while the latter was not. To see if this was the correct explanation we repeated the experiment with each of the conditions 1-3 in separate blocks, using a separate sample of control subjects (7 females, matched by age with the original control group.). Once again, we found that search times in Condition 3 were faster than in Condition 1 ($p<0.05$ for all subjects). The deleterious effect of irrelevant colour is thus not due to blocking with the colour-relevant task, and probably has the same explanation as the effect reported by Morgan et al. (1992).

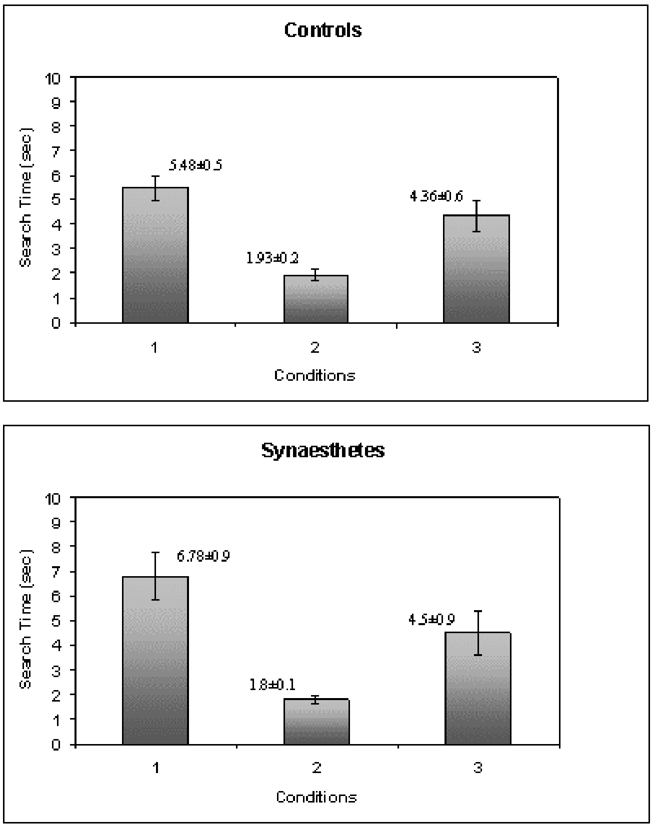


Fig. 5-5: Results for Experiment 1: Bars represent the mean of mean reaction time for all observers, including control and synaesthetes. Condition 1: Look for the line with unique tilt among coloured lines. Condition 2: Look for the line with unique colour among coloured lines. Condition 3: Look for the line with unique tilt among black lines. The two groups show the same effect: looking for the unique coloured line was the easiest task. Looking for a unique tilt was faster when surrounding lines were all black compare to when they were differently coloured.

In summary, this first experiment establishes that search for a graphemically-unique target is improved if it is uniquely coloured as well; and impaired if irrelevant colour is added to the graphemes relative to search when all the graphemes are black. These results set the stage for the second experiment, which tests whether similar effects will be found with synaesthetic colours.

5.6.2 Experiment 2

Percent correct were high for all subjects in all three conditions and did not differ significantly between conditions. Reaction time data are shown in Fig.5-6. Our predictions were that synaesthetes would be faster than controls in the unique condition, but slower in the non-unique one. However, as is evident from inspection, there was no significant difference between groups in either the two conditions. (F = 4.25, p = 0.87), and there were no significant interactions (F = 4.25, p = 0.56) between groups and conditions.

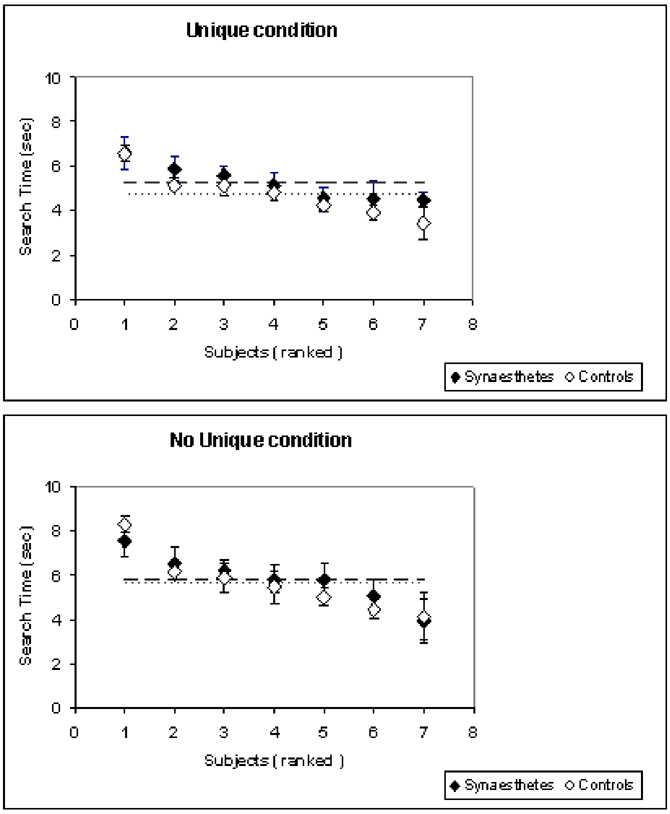


Fig. 5-6: Results for Experiment 2. The graphs show the mean Reaction Time for each subject in the two conditions. The two lines represents the means of the two group: the black dashed one for synaesthetes, the dotted one for controls. There was no significant difference between synaesthetes and control.

A further hypothesis was that synaesthetes would be better in the unique condition compared to the non-unique one, but there was no significant difference (F = 4.25, p = 0.11). Individual t-tests showed a significant difference between conditions in 2 of the 7 synaesthetes, but this was also found in 5 of the controls (see Fig. 5-7). The probable explanation of this difference is that a smaller range of numbers was used in the unique vs. the non-unique condition, as explained in the Methods section.

	SYNAESTHETES						
	CB	EB	JG	JL	LW	SK	VL
Non-unique Vs unique	t = 1.59 p = .11	t = 0.8 p = .42	t =-2.39 p< .05	t =1.32 p = .18	t =1.61 p = .1	t =0.75 p = .45	t= 3.28 p< .05

	CONTROLS						
	ER	JM	HT	JE(51)	CW	KB	JE(46)
Non-unique Vs unique	t =2.27 p < .05	t = 2.9 p < .05	t = .47 p = .63	t = 2.22 p <.05	t = 3.44 p < .05	t =1.45 p =.14	t =2.51 p < .05

Fig. 5-7: T-test values for Experiment 2: Individual t-tests showed a significant difference between conditions in 2 of the 7 synaesthetes, but this was also found in 5 of the controls. $\alpha = 0.05$

5.7 Discussion

Our data have failed to confirm the prediction that synaesthetes will be impaired by synaesthetic colours in a visual search task where real colours impair performance. They therefore do not support the proposition that synaesthetic colours, at least in our subjects and with our experimental design, are ‘perceptual’ rather than ‘conceptual’. This cannot be explained by the test being insufficiently sensitive, since Experiment 1 showed that synaesthetes and controls were impaired by ‘real’ colours. The possibility remains that synaesthetic colours are ‘perceptual’ but too weak to impair performance like ‘real’ colours. Against this, strong effects of synaesthetic colour have been reported in previous experiments. In the Kim, Blake, & Palmeri (2006) study of apparent motion, the effect of synaesthetic colour was equivalent to real colour in one subject, and stronger in the other.

On the other hand, our results are concordant with others suggesting that synaesthesia may result from the activity of areas concerned with language and visual feature segregation or that colours perceived might arise from frontal brain areas that lie beyond the perceptual processing hierarchy (Paulesu, 1995, Rich, 2002, Schiltz, 1999).

It has been suggested that only ‘lower’ rather than ‘higher’ synaesthetes, or alternatively, only ‘projectors’ experience synaesthetic colours that can help them in visual search (see Introduction). We do not know into which category our synaesthetes fell. We concede the possibility that our sample were of the kind, for whom subjective colour is not an aid to visual search. On the other hand, our synaesthetic subjects had been previously characterised in the fMRI study by Nunn et al. (2002) showing activation of V4/V8 by spoken words having stable colour associations, and this would seem to indicate a relatively low level colour signal. Either the V4/V8 signal is not sufficient to help with visual search, or our subjects have V4/V8 activation to spoken but not visual input. Our result, taken with that of Nunn et al. (2002) have the interesting implication that V4/V8 activation is not necessarily associated with the perceptual experience of colour.

Chapter 6

Conclusions

6.1 Summary of Arguments and Findings

The original purpose of the investigations described in this thesis was to see whether the subjective colours experienced by grapheme-colour synaesthetes could make targets more ‘salient’ in visual search and in crowding tasks. The sample of synaesthetes used in the experiment had been previously used in an fMRI study by Nunn et al. (2002) showing that there was V4 activation in these subjects by auditory input of digits. As far as this aim is concerned, the results were entirely negative. We were unable to confirm previous reports that visual search for an unique shape was aided when that shape had an unique synaesthetic colour. An investigation of colour camouflage by synaesthetic colours, designed to make synaesthetes perform worse than normal controls, was equally unsuccessful. It is possible that we used the ‘wrong kind of synaesthete’ for our experiments, or that the task was not ideal for demonstrating an effect. On the other hand, the literature presented in Chapter 5 has shown that previous studies claiming a perceptual reality of synaesthetic colours have largely failed to control for motivation and experimenter expectancies. An increasing number of studies, of which ours is just one, are beginning to question the existence of pre-attentional effects of synaesthetic colours. The jury must therefore be regarded as still ‘out’ on the question of the perceptual reality of synaesthetic colour.

Given the failure of the experimental study of synaesthetic colour, we switched the focus of the investigation to crowding in visual search in normal subjects. A review of the literature showed that stimulus ‘salience’ has been a key concept in both areas. In visual search experiments, salience has been defined by the phenomenon of ‘pop out’, in which search time for a target is independent of the number of distractors. In ‘crowding’, salience has been invoked to explain why a target can at least partially escape crowding by differing from distractors in depth or colour. Hitherto, however, there has been no demonstration that there is a connection between ‘pop out’ and relief from crowding using the same stimulus configurations. Using the dimensions of motion direction, colour and temporal frequency, we have shown for the first time that there is a correlation between visual search times and crowding. Stimulus salience therefore emerges as a key concept crossing the boundary between the visual search and the crowding literature.

Recent explanations of crowding have invoked the idea that crowding occurs when there is an inappropriate integration of feature information across objects, at the expense of feature binding into objects. One version of this idea is that there is unavoidable, if undesired, integration of features into texture. A quite different idea is that crowding increases early noise in feature representation, so that target and distractor are more likely to be confused. Consider a crowded display consisting of a set of vertically-oriented distractors and a single non-vertical target that has to be judged ‘clockwise’ or ‘anticlockwise’. According to the averaging theory the observer will report the noisy average of the ensemble. According to the early noise theory, supplemented with the ‘Max’ rule of Signal Detection Theory, the observer reports the orientation of the stimulus seeming to have the greatest absolute deviation from the vertical, which may be the target, but which may equally be a distractor. If, instead of reporting the direction of offset of the stimulus, the observer is required to match its perceived orientation to that of an adjustable probe, the two theories make quite different predictions. On the averaging model, the distribution of matches will be unimodal, centred on the vertical; on the ‘Max Model’ it will be bimodal, with a minimum at the vertical.

In The following Appendix, I report an experimental test of these different predictions of the averaging vs uncertainty models. Using the same display parameters of Parkes et al (2001)I was able to confirm their finding that threshold fell linearly with the number of ‘distractors’ that had the same tilt as the target. This is in conformity with the averaging model. However, the distribution of matches in the matching experiment was bimodal, in agreement with the non-linear ‘Max’ rule. Furthermore, observers were able to localise the target with above-chance accuracy, contrary to both the findings and predictions of Parkes et al. (2001). On the other hand, localisation accuracy was not as high as that predicted by the

‘Max’ model.

To conclude, based on the results of this study, we propose that uncertainty contributes to the sensitivity impairment in the presence of flankers to a peripheral target, but it is not the sole cause. In our crowding task what may be happening is that the system cannot avoid supporting its decision on a channel representing a flanker, or group of flankers, if it produces a response stronger than that representing the target. What remains to be determined is (i) the size of the regions in which orientation estimates are averaged, (ii) the number of these regions that influence observer responses and (iii) the exact form of the decision rule. In any case, we have enough evidence to suggest that identification of a crowded target’s tilt is not limited to an array’s global average orientation. Some form of intrinsic uncertainty plays a role.

Appendix

Non-linear integration of crowded orientation signals

Crowding is a condition in which identification of a visual target is impaired by abutting elements. One leading account of crowding of individual features such as orientation explains this phenomenon as compulsory averaging of the signals from target and flankers before the site of perceptual awareness (Parkes et al., 2001). Recent measures of internal response distributions with sparse, uncrowded stimuli but an identical task (Baldassi et al., 2006) have shown that the observers’ internal representation of stimulus tilt arises from non-linear combination of independent noisy responses followed by a maximum-of-outputs rule (Baldassi and Verghese, 2002). In the present study, we have measured internal distributions under the crowding conditions of Parkes et al. (2001) and found bimodality, which is the signature of non-linear combination of information from each stimulus. In addition, we asked our observers to *localize* a single flanker tilted like the target. Increasing the tilt of the target caused *improvement* of localization performance, but not as much as predicted by a strict maximum-of-outputs rule. Our results challenge the notion of global averaging and suggest at least some conscious access to local estimates of orientation.

A.1 Introduction

Identifying a target in a cluttered visual scene can be a very difficult task. Moving the stimulus to the periphery of the visual field, decreasing the size of the elements and the distance among them makes such a task even harder. This effect is called crowding.

In a study that is the directly connected to our present investigations, Parkes et al. (2001) have studied the effect of crowding on simple feature processing, i.e. orientation discrimination. They measured the effect on thresholds of flanking a small tilted target in the periphery with a number of oriented elements all displaced within the spatial range of crowding. Keeping the overall number of elements fixed, they increased gradually the number of flankers that were tilted like the target and observed that thresholds were reduced linearly (on log-log scales) with increasing number of tilted flankers even though observers knew they had to judge only the target that was sitting centrally on the array of Gabor patches (see Figure A-1). In a separate task, they tilted three of the nine elements according to a vertical vs. a horizontal configural arrangement and observers were asked to identify the orientation of the configuration, a task implying that observers localized individual patches. This task was very difficult to perform. The results of this study have led to the suggestion that crowding of individual features, such as orientation, can be explained as a compulsory averaging of information from targets and flankers before the site of conscious evaluation of the target tilt (Parkes et al., 2001). In its strictest form, this model assumes that the system averages orientation estimates for all targets and flankers. Since the number of elements was fixed to nine, the overall amount of noise (i.e. the standard deviation of the function assuming gaussian distributions) is unaffected, while the mean of the distribution increases linearly with the number of positive signals introduced. A fundamental consequence of this model is that, within the range of crowding, it would be impossible to segregate individual visual targets by having access to their information in isolation. A similar averaging rule has been suggested in the domain of textural integration and visual search (Dakin and Watt, 1997, Baldassi and Burr, 2000, Morgan et al., 1998).

In the domain of visual search a very similar effect on orientation discrimination thresholds is predicted by a version of the Max of Outputs rule (Palmer, 1994; Palmer et al., 1993) developed for ‘two-tailed’ orientation discrimination tasks, in which the target randomly assumes one of two possible values around a reference (e.g. CW or CCW tilts away from vertical). It assumes that each stimulus in the array is monitored by at least two detectors (one for each direction of tilt), whose output is noisy. There is no further perceptual processing of the display (such as averaging or summation) and the decision rule is based on the preferred direction of the detector yielding the *strongest* response. Because of noise, there are two main consequences of adding distractors (or flankers) to the display. One is that the more distractors are included in the computation, the more it will be likely that one of the distractors will be signalled as more tilted than the distractors, in both the same and the opposite direction of tilt. In the latter case the observer produces an error. The other is that increasing the number of elements causes the maximum response from each of the two classes of detectors monitored to increase, and this is reflected in the response distributions as a lack of responses at small tilts and an increase of response at high tilts. In other words, the distribution of maxima from the two oppositely tilted detectors monitored in the task becomes increasingly bimodal when the number of elements increases. Baldassi & Verghese (2002) have shown that this model quantitatively explains both the thresholds and the change in shape of the psychometric function with set size better than an Averaging model, and Baldassi & Burr (2004) have shown that it fits well with the observed pop-out of a luminance target at threshold and with the flattening of the set size function at high set sizes. Using a novel psychophysical technique that probes the internal stimulus representation directly, Baldassi et al. (2006) were able to demonstrate that in similar search displays the distribution of reported tilts follows strictly the predictions of the Signed Max model, that is unimodal at set size 1 and increasingly bimodal at physical or cue-defined set sizes higher than 1. Interestingly, the use of external noise in the set size 1 condition to match the sensitivity of set size 8 changed the spread (i.e. the standard deviation) but not the shape of the distribution, which remained unimodal.

The present study bears on the speculation that if crowding of oriented signals relies on compulsory averaging of noisy information this should have a clear signature in the measure of internal response distributions, in that the distribution should broaden when flankers are displayed, with no change in shape. On the contrary, if the limits come from integration rules different from Averaging, this should reflect into a change of shape of the distributions with changing the amount of flankers around the target.

We did not have specific reasons to commit in advance to a Max or to any other non-linear integration rule for the crowding task. However, using a number of converging measures, we show that the orientation signal carried by a crowded target is combined in a non-linear fashion and departs in many ways from the predictions of an Averaging model, showing the signature of some form of uncertainty over the target identity combined with an inefficient use of detectors at very small scales.

A.2 General Methods

A.2.1 Stimuli

The parameters of the stimuli used in our study were designed to reproduce exactly the display used by Parkes et al., 2001. Stimuli were generated in Matlab, using Psychophysics Toolbox extensions for Macintosh (Brainard, 1997, Pelli, 1997) and presented with a Mac G3 computer on a 17” Sony display at 75 Hz refresh rate. The individual elements were Gabor patches (12 c/deg sinusoidal gratings of 90% contrast, 50% for the noise experiment, and 29 cd/m2 mean luminance, windowed within a circular Gaussian aperture of s=0.083° space constant), at 2.5º eccentricity. A stimulus set comprised 1 central target that was always tilted, surrounded by 8 elements, all vertical. The centre-to-centre separation of the central target with each flanker was equal to $\lambda 3\sqrt{2}$. In the intrinsic uncertainty control experiment the Gabor patches were surrounded by a white annulus of 62 cd/m2. Target tilt was selected at random from 12 different orientations, ranging between ±35°. In the noise experiment, the tilt of the target homed in around threshold based on the QUEST algorithm (Watson and Pelli, 1983). The stimuli were displayed for 100 ms, followed by a mean grey page for 200 ms, and then the response page appeared.

A.2.2 Observers and procedure

Eleven observers participated in the experiment. They had normal or corrected to normal vision and they were all naive to the goal of the study, except for one who was the author (CG). Observers collected data for different set sizes in different blocks. The overall number of trials, executed in blocks of variable length (50 to 120) was about 600, with a variability depending on the stability of the results and the particular condition. Observer CB performed many more trials (~1800) as he participated in the pilot phase of the experiment.

In the magnitude matching task, observers were presented with 12 response probes whose tilts corresponded to the stimulus set from which the target was sampled. In the magnitude estimation task the response page was a probe similar to the target (but twice larger) that could be rotated by lateral motion of the mouse. Immediately after the response a blank page of 400 ms was displayed and the following trial started automatically. No feedback of any kind was provided.

In both tasks, we expected that the distribution of reported tilts by the observers to probe the distributions of internal noisy states of representation of the tilt, thus enabling to assess more directly the mechanism of crowding of orientation information.

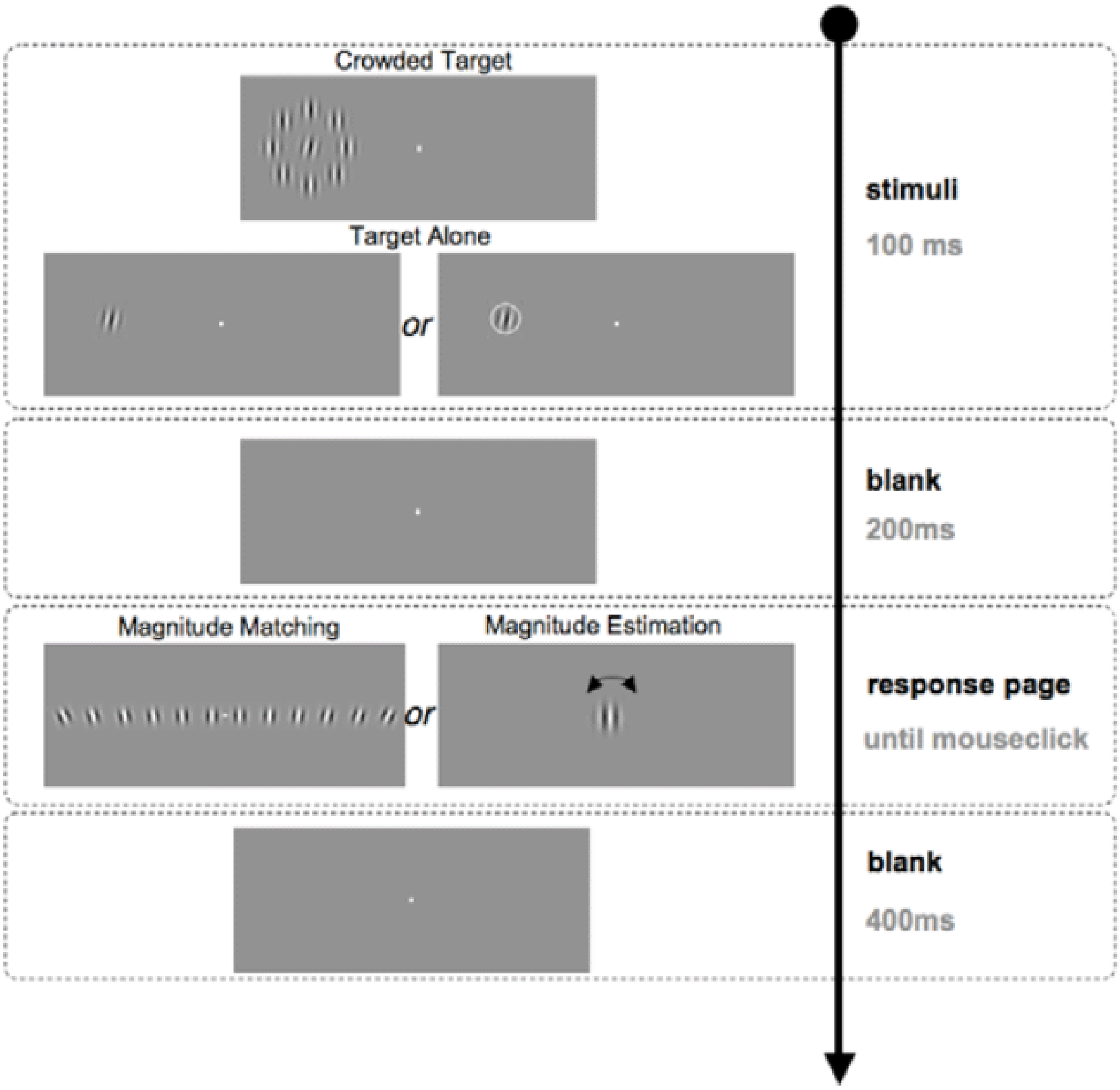


Fig. A-1: Experimental sequence and stimuli. The stimuli matched exactly the parameters of Parkes et al. (2001) but were shown according to a vertical, rather than horizontal reference, After the stimulus display, a response page was shown requiring one of two alternative tasks. In the magnitude matching task, observers were presented with 12 response probes whose tilt corresponded to the stimulus set from which the target tilt was sampled. In the magnitude estimation task the response page consisted of a probe resembling the target (but 2 times larger) that could be rotated by lateral motion of the mouse. Immediately after the response a blank page of 400 ms was displayed and the following trial started automatically. No feedback of any kind was given. One of the observers was tested in the target alone condition with an outlining circle to reduce intrinsic uncertainty.

A.3 Results

A.3.1 Thresholds vs Number of Targets

Parkes et al. (2001) used a simple binary task to show that increasing the number of tilted patches carrying the same tilt helped performance following a slope of 1 on log-log coordinates. This result implies crowding, in fact the surrounding elements influence thresholds even though the observers knew they have to judge only the tilt of the central element, and was explained as compulsory averaging of orientation signals from target and flankers. In order to compare the outcome of our measures of response distribution to the results shown in Parkes et al. (2001), we first verified Parkes et al.'s results with our magnitude matching technique.

We asked two naïve observers, CB and JC, to indicate the perceived direction and magnitude of tilt of the central target by clicking on one of the 12 response probes appearing after the stimulus. The number of tilted targets was randomly interleaved for observer CB while it was blocked for JC. Observer CB ran many trials (about 1800 for each condition), so we could use his data also for the analysis of the response distribution shown later, while JC ran only about 180 trials per condition, enough to measure thresholds but not for sketching the response distributions.

We analyzed the data by classifying clockwise vs. counterclockwise responses in a standard binary fashion to produce psychometric functions that were fit with a cumulative gaussian function to estimate thresholds, corresponding to the level of tilt producing 75% of correct tilt direction judgements.

Figure A.2 shows the orientation discrimination thresholds for the two observers for judging the direction of tilt of a central Gabor patch surrounded by an array of 8 flankers, only a (variable) proportion of which carried the same signal of the target.

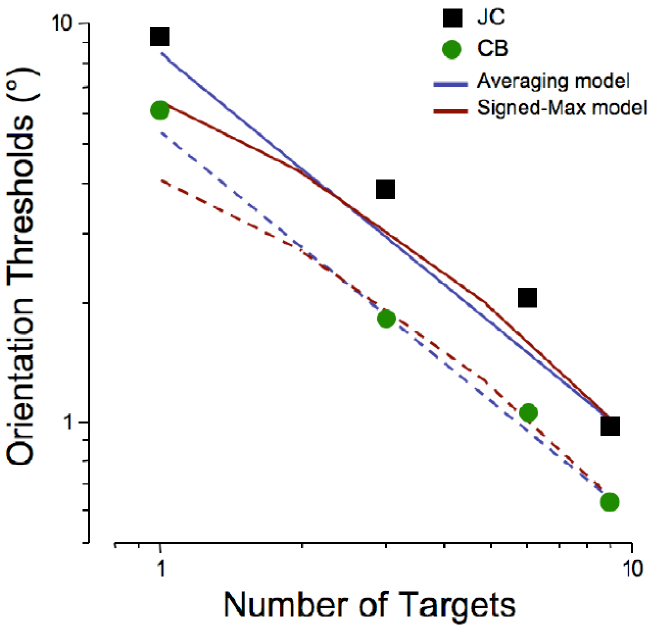


Fig. A-2: Threshold estimation from the magnitude matching procedure. The graph shows orientation thresholds as a function of the number of targets, in log-log axes. The symbols represent the measures for the two observers, while the blue and purple lines are the predictions of the Averaging and of the SignedMax model, respectively. Orientation sensitivity (reciprocal of threshold) increases with the number of patches having the same tilt as the target . The rise of performance on log-log axes is close to the predictions of both models for all conditions but the one with only the central target, that is well predicted by the Averaging model while the Signed-Max model predicts a better performance than the data. These data are a very good replication of those by Parkes et al. (2001), suggesting that the magnitude matching task does not alter the decision strategy in any way.

These results confirm and replicate the results of Parkes et al. (2001). In particular, the more tilted elements there were in place of vertical flankers, the better were the thresholds, implying an increase of orientation sensitivity, dropping to 1° for JC and 0.65° for CB. The rise of performance was very well predicted by the Averaging model (purple line), as shown by Parkes et al. (2001), even though the Signed-Max model did as well when the number of targets was higher than 1. When only one target is present the latter model predicts a better performance than the data, consistently for both observers. In any case, the good match between our and Parkes et al.'s results legitimates the use of the magnitude matching technique to further characterize the mechanisms underlying crowding of oriented signals, in that it leaves the pattern of thresholds unaffected relative to standard, binary psychophysical tasks.

A.3.2 Matching and estimation of target tilt

A.3.2.1 Measuring the target alone

In the second stage of our study, we measured the response distributions underlying crowding of oriented targets by comparing the responses on the orientation of a small Gabor target under different crowding regimes. We asked three observers to classify both the tilt direction and the tilt magnitude of a target appearing 2.5° to the left or to the right of fixation either in isolation, or in the presence of eight vertical flankers. This condition was used as baseline for both the measurement of the crowding effect and that of the response distribution. In fact the measure of the target displayed in isolation should reveal a response distribution of gaussian shape whose parameters depend strictly on the physical signal (i.e. different tilt offsets from vertical) and the noise (internal and external). The second condition, with eight neutral flankers, was the one showing the strongest impact of crowding in Parkes et al. (2001) study as well as in our replication described in the previous paragraph.

CB was tested with the magnitude matching paradigm, while GB and CG used the magnitude estimation paradigm. GB was also tested in a different condition with an outlining circle around each patch (described later). Responses were scored correct if the sign of the tilt was correctly identified, irrespective of the magnitude match. For each observer and condition, the identification responses were binned into three classes of discriminability: near-threshold (67 to 83% correct responses), below-threshold (less than 67 % correct), and above-threshold (greater than 83% correct). Only distributions for near-thresholds angles are presented here as they provide the best information about the shape of the response distributions, free from floor or ceiling effects (Baldassi et al., 2006).

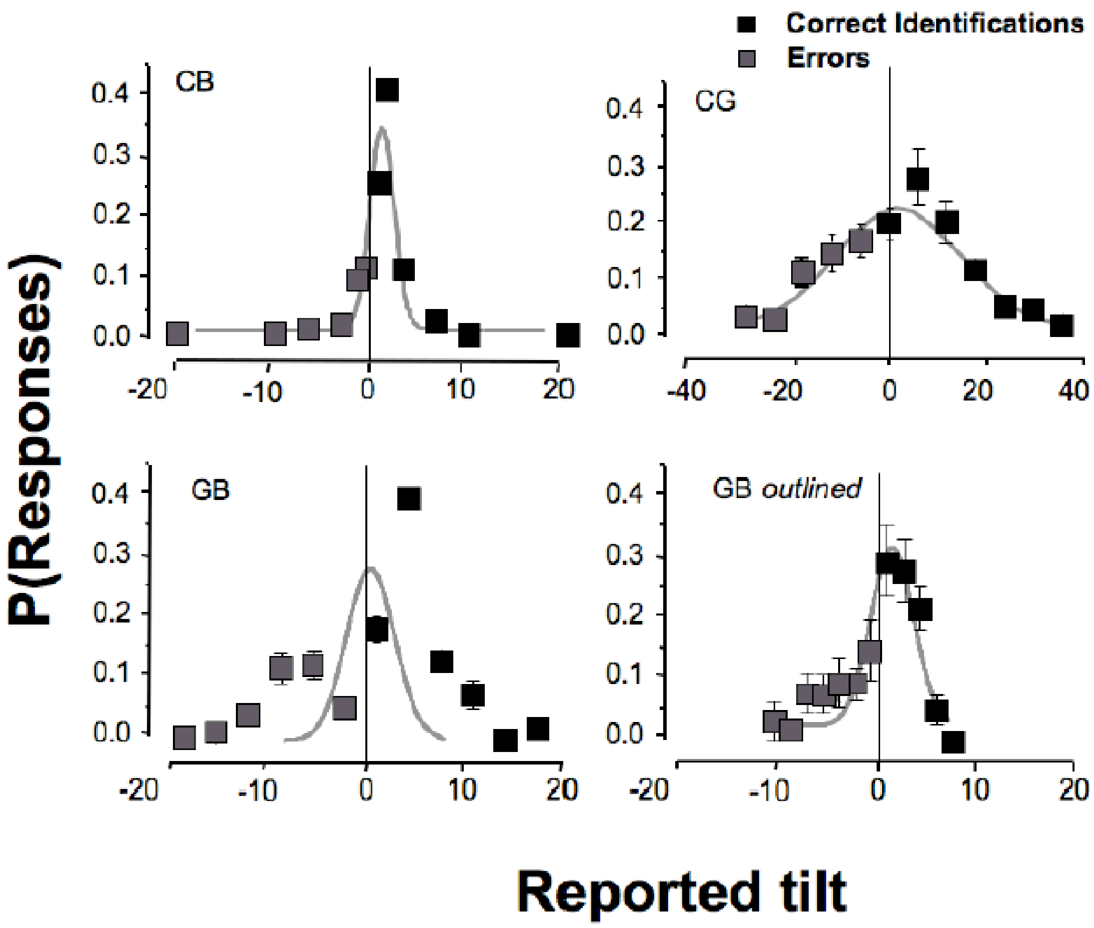


Fig. A-3: Response distributions for the target alone condition. Each panel plots the proportions of responses as a function of the reported tilt for three observers (GB was tested in two conditions, see text). Grey squares, to the left of 0, represent errors, while black squares represented correctly identified tilt directions. As predicted by both models two out of three observers show unimodal distribution well approximated by gaussian functions. Observer GB does not follow this pattern in the basic condition but does it when the target location is clearly outlined by a circle around the target.

Figure A-3 shows results for all observers in the target alone condition. The symbols show a histogram of the proportion of reported tilts. Grey squares represent errors and black squares correct identifications. The error bars represent the standard error of the mean (S.E.M.) calculated by bootstrap (Efron, 1993). When not shown, the S.E.M. is smaller than the symbol.

All observers except one (GB), show unimodal, gaussian-like distributions as predicted by both models. The solid line represents the best fitting gaussian function underlying the distributions. The spread of these distributions was in all cases (except GB's in the basic condition) in good agreement with the threshold at set size 1 for each observer, consistent with our assumption that the distributions we measured corresponded to the internal states generated by our stimulus as predicted by the Signal Detection Theory (Green and Swets, 1966). The criterion to decide whether a distribution was unimodal or bimodal was the following

(Baldassi et al., 2006). The largest positive and negative responses were selected as potential peaks. If any data points between them were significantly lower than both these peaks (bootstrap t test, $p < 0.01$) then the distribution was classified as bimodal. All the curves of the Target Alone condition, except for GB's with the standard display, were classified as unimodal.

The bottom right panel of Figure A-3 shows GB's data in a slightly modified version of the task. In fact, we wondered if the bimodality shown in this observer could be attributed simply to a fault of the estimation task measurement when observers had to report small tilts, or to a more subtle effect like the intrinsic uncertainty (Pelli, 1985) that could be generated by a very small, isolated peripheral patch of high spatial frequency even if it was displayed at high contrast. Indeed, even if the target had high contrast, observers sometimes reported not to see it clearly (or at all). To answer this question we re-ran this condition in observer GB with an outlining cue that was clearly cueing the location of the stimulus, possibly reducing intrinsic uncertainty. The bottom right panel of Figure A-3 shows the response distribution of this observer with an outlined patch showing a return to a gaussian, unimodal shape of the distribution, coherent with a reduction in intrinsic uncertainty.

A.3.2.2 Measuring the crowded condition

Results for the crowded condition are shown in Figure A-4. Again, proportions of responses for three observers are plotted as a function of the reported tilt. The bottom right panel reports the distributions of GB with outlining circles around both target and flankers; we reasoned that in the crowded condition the flankers should reduce the positional uncertainty about the target location, but this measure was performed to control for potential artefacts from the outlining circles. In this case the pattern of distributions changes drastically compared with the Target Alone condition. In all cases the criterion for bimodality was met, with larger tilts preferred to smaller tilts for both errors and correct identifications.

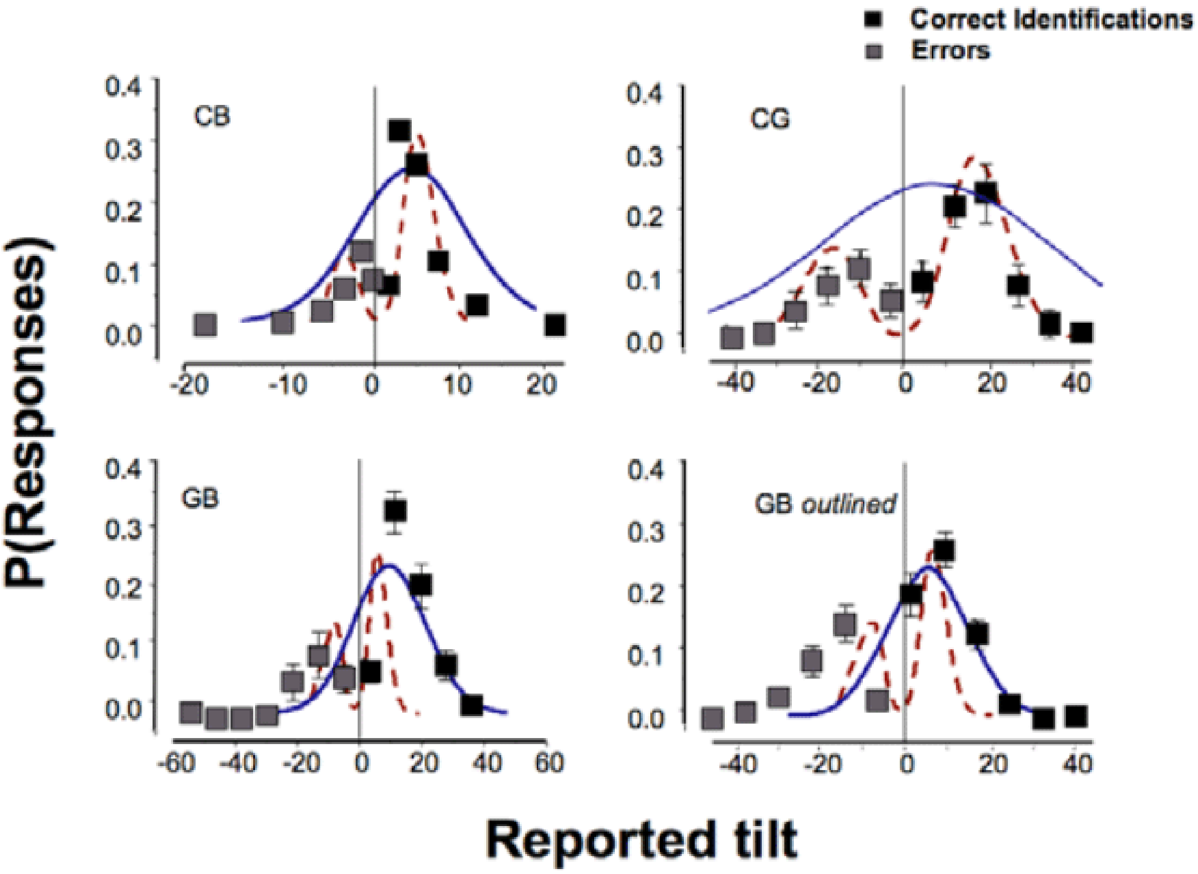


Fig. A-4: Response distributions of the crowded condition. Each panel plots the proportions of responses as a function of the reported tilt for three observers (GB was tested also with outlined patches. Symbols are like those of Fig. A-3: grey squares, to the left of 0, represent errors, while black squares represented correctly identified tilt directions. The solid blue line represents simulation for the Averaging model, while the dashed red line represent simulation for the Signed-Max model. All observers exhibit a low rate of responses for small tilts and an increase for larger tilts implying bimodality, but the tails of the empirical distributions are generally larger than the predictions of the Signed-Max model possibly implying that the flanker introduced additional sources of noise.

This pattern was very close to that obtained in the uncrowded condition of visual search (Baldassi et al., 2006). We therefore modelled the data simulating both the Averaging model (solid blue lines, Fig. A-4) and the Signed-Max model (dashed red lines, Fig. A-4). The only parameter used was the internal noise for one element (i.e., the threshold at set size 1), that was successful in describing the response distributions at set size 1. The MonteCarlo simulations rule out the strict Averaging model, in that it predicts unimodal distributions independently of whether or not the flankers were there. The Signed-Max model was a very good fit of the data in two out of three observers. In the two conditions tested in observer GB, even though the bimodality is clear (and significant based on the bootstrap test), the prediction of the Signed-Max model, represented by red dashed lines, underestimates the width of the empirical function, possibly indicating some additional source of noise in the computation, that is often found in set size modulations (Wilken and Ma, 2004) or the use of a hybrid model, in which the max rule is applied to regional rather than global orientation estimates (see below).

A.3.3 Locating individual crowded stimuli

The data shown so far seem to support an uncertainty explanation of this sort of crowding. However, we still cannot completely rule out that this pattern was due to a version of the Averaging model that combines outputs from independent, first stage filters after a non linear transformation, such as squaring (Heeger, 1992). If the output from each local element is raised to a power greater than 1 before the averaging stage, then the element with the strongest output will count more and more with increasing exponent. The response distribution generated by a similar averaging rule will lose linearity, approximating the one predicted by the Signed-Max rule, showing a bimodality that sharpens with higher exponents.

We then decided to test if observers would be able to *locate* the target. The paradigm was slightly varied so that one of the flankers was tilted like the target, and observers were required both to locate it and to estimate the magnitude of its tilt. This second task was performed to check that the modifications introduced in the task did not alter the observers' strategy when judging the tilt of the targets relative to our previous experiments. Note that in this condition there are *two* targets, the central one and one randomly selected from the crown of the 8 flankers.

The top panel of figure A-5a shows the proportion of correct localization, independent of the identification response, at different identification performance levels relative to threshold (BT=below threshold; NT=near threshold; AT=above threshold). The mean of our four observers (standard errors are smaller than the symbols) clearly show a consistent and significant pattern of *improvement* of localization performance with increasing discriminability, that departed from guessing (coinciding with the predictions of the Averaging model, the blue solid line), but not as efficient as predicted by an uncertainty model (purple dashed line) that selects the location with the strongest output. The bottom panel of figure A5b

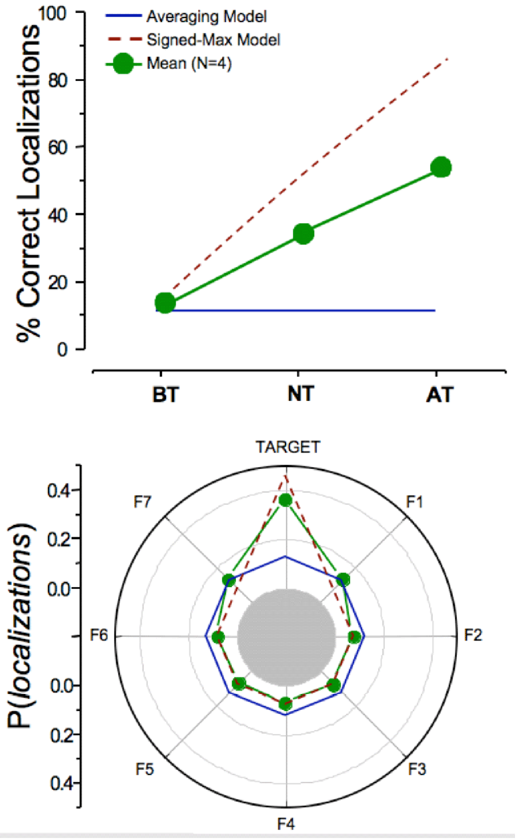


Fig. A-5: Results for the location task. The top panel plots percent of correct localizations as a function of different target tilts leading to different identification performances (BT=below threshold, NT=near threshold, AT=above thresholds). Localization performance, represented by the green circles, improves significantly with target tilt, against the Averaging prediction (blue solid line) but less rapidly than the predictions of an uncertainty model (purple dashed line) that locates the target based on the strongest output. The bottom panel shows the proportion of responses for each location when the target tilt was correctly identified. Note that the axis goes to negative values, corresponding to the central grey circle, an artefact we have introduced to increase the visibility. Polar angle reflects the selected position, relative to the tilted flanker position, that in the plot is normalized to its North location. Flankers are numbered in a clockwise, progressive order (F1 to F7). The green circles show the mean of four observers and its standard error (smaller than the symbols). The blue solid line shows the guessing rate, coincident with the predictions of the Averaging model, while the purple dashed line the prediction for the Signed-Max model. The position of the target was reported significantly more often ($p<0.01$) than the other locations for correct identifications and the locations adjacent to the target were reported more often than those farther away (Tukey test α=0.05).

shows a polar representation of the selected locations relative to the target when the target tilt was correctly identified. The tilted flanker location was normalized to be plotted at the North location, and each green circle is the mean of the four observers (with SEM). The blue solid line shows the pattern of random localization that coincides with the predictions of the Averaging model, while the purple dashed line shows the predictions of the Signed-Max model, represented by a peak of localization responses at the target position and the other seven locations were randomly selected. When observers identified the tilt of the two targets, they could also locate the tilted flanker significantly more often than the other locations (ANOVA test, p<0.01, F= 79.3628, Tukey test < = 0.05), ruling out the Averaging model. The only difference between the data and the predictions of the Signed-Max model was in the 'nearby flanker' errors shown in the data.

It thus appears that observers are not limited to a global average orientation, and are able to locate tilted targets in crowded arrays with some degree of accuracy. These localization results diverge from those obtained by Parkes et al. (2001), who found that observers were unable to identify the configuration of three tilted elements in a crowded array. One possible reason for our divergent results is our size of target angles. Parkes et al. do not report the size of theirs, only that performances, "did not improve with target tilt". When our targets were Below Threshold, our observers' localizations were no better than chance. Perhaps Parkes et al's target tilts were similarly small.

A.4 Discussion and Conclusions

In this study we investigated the mechanism underlying crowding of orientation signals, previously explained by compulsory averaging before awareness. The study by Parkes et al. (2001), to which our work is directly connected, reported a linear decrease of thresholds with increasing number of targets from an array of nine oriented signals. In our study, we replicated the effect but increased the scope of the results by concurrently measuring the distributions of perceived magnitude. The distributions showed bimodality of perceived tilts, which is a signature of non-linear combination of information from the elements composing a crowded display.

For further evidence of conscious access to local orientation signals, we tested the observers' locating ability. Whereas location would be impossible without some such access, performances fell well below the Max-rule predictions (see Fig A-5). Perhaps observers used some hybrid strategy, averaging orientation estimates in different regions of the stimulus array, and then applying a max rule to those averages. Another possibility is that, when identifying target tilt, observers use some non-linear combination of tilt estimates (e.g. the max rule); but their localization responses are based on texture borders. Targets form a texture border with each adjacent distractor. On trials in which only one of these borders can be detected, observers may select the adjacent distractor, rather than the target's true location (Solomon and Morgan, 2001). Mislocalization to adjacent elements can be associated to the illusory conjunction effect, the condition in which the observer combines different characteristics of nearby elements susceptible to spatial crowding (Treisman and Schmidt, 1982), or to the similar phenomenon of feature inheritance (Herzog, 2001).

To conclude, based on the results of this study, we propose that uncertainty contributes to the sensitivity impairment in the presence of flankers to a peripheral target, but it is not the sole cause.

In our future studies we would attempt to investigate whether crowding is the result of the visual system limitation (e.g. the system cannot avoid basing its decision on a channel representing a flanker, or group of flankers, if pruducing a response stronger than that representing the target). Many other researches can also follow this initial results. For example is still to be determined either the size of the regions in which orientation estimates are averaged or the number of these regions that influence observer responses. Very interesting would be to evaluate the exact form of the decision rule that acts in these specific conditions.

References and Bibliography

Amir, Y., Harel, M., & Malach, R. (1983). Cortical hierarchy reflected in the organization of intrinsic connections in macaque monkey visual cortex. *Journal of Comparative Neurology,* **334**.

Andriessen, J. J. and Bouma, H. (1976). Eccentric vision: adverse interactions between line segments. *Vision Research,* **16,** 71-78.

Anstis, S. M. (1974). A chart demonstrating variations in acuity with retinal position *Vision Research,* **14,** 589-592.

Aubert, H., Förster,R. F. (1857) *Beiträge zur Kenntnis des Sehens.*, Berlin.

Baldassi, S. and Burr, D. C. (2000). Feature-based integration of orientation signals in visual search *Vision Research,* **40,** 1293-1300.

Baldassi, S. and Burr, D. C. (2004). "Pop-out" of targets modulated in luminance or colour: the effect of intrinsic and extrinsic uncertainty *Vision Research,* **44,** 1227-1233.

Baldassi, S., Megna, N. and Burr, D. C. (2006). Visual clutter causes high magnitude errors *PLOS Biology,* **4,** e56.

Baldassi, S. and Verghese, P. (2002). Comparing integration rules in visual search *Journal of Vision,* **2,** 559-570.

Baron-Cohen, S., Burt, L., Smith-Laittan, F., Harrison, J. and P.Bolton (1996). Synaesthesia: prevalence e familiarity. *Perception,* **25,** 1073-1079.

Baron-Cohen, S., Burt L.,Smith-Laittan F.,Harrison J.,Bolton P. (1996). Synaesthesia: Prevalence and familiarity. *Perception,* **25,** 1073–1079.

Baron-Cohen, S., Harrison J.,Goldstein, L. H.,Wyke, M. (1993). Coloured speech perception: Is synaesthesia what happens when modularity breaks down? *Perception,* **22,** 419–426.

Baron-Cohen, S., Wkye, M., Binnie, C. (1987). Hearing words and seeing colours: an experimental investigation of a case of synaesthesia *Perception,* **16,** 761-767.

Beck, J. (1982) In *Organisation and representation in perception*(Ed, Beck, J.) Lawrence-Earlbaum Associates, Inc., Hillsdale,N.J.

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

Blasdel, G. G., Lund, J. S., & Fitzpatrick, D. (1985). Intrinsic connections of macaque striate cortex: axonal projections of cells outside lamina 4C. *Journal of Neuroscience,* **5,** 3350-69.

Bouma, H. (1970). Interaction effects in parafoveal letter recognition *Nature,* **226,** 177-178.

Bouma, H. (1973). Visual interference in the parafoveal recognition of initial and final letters of words. *Vision Research,* **13,** 767-82.

Bouma, H. and Andriessen, J. J. (1970). Induced changes in the perceived orientation of line segments *Vision Res,* **10,** 333-49.

Brainard, D. H. (1997). The Psychophysics Toolbox *Spatial Vision,* **10,** 433-436.

Butler, T. W. and Westheimer, G. (1978). Interference with stereoscopic acuity: spatial, temporal, and disparity tuning. *Vision Research,* **18,** 1335-1342.

Callaghan, T. (1984). Dimensional interaction of hue and brightness in preattentive field segregation *Percept Psychophys,* **36,** 25-34.

Callaghan, T. C., Lasaga, M.I., Garner, W.R. (1986). Visual texture segregation based on orientation and hue *Percept Psychophys.,* **39,** 32-8.

Carter, R. C. (1982). Visual search with colour *Journal of Esperimental psychology: Human Perception and Performance,* **8,** 127-136.

Casco, C., Campana, G., Grieco, A., Fuggetta, G. (2004). Perceptual learning modulates electrophysiological and psychophysical response to visual texture segmentation in humans. *Neurosci Lett.,* **371,** 18-23.

Cass, J. R., & Spehar, B. (2005). Dynamics of collinear facilitation are consistent with long-range horizontal striate transmission. *Vision Research,* **45,** 2728–2739.

Cave, K. R. and Wolfe, J. M. (1990). Modeling the role of parallel processing in visual search *Cognitive Psychology,* **22,** 225-271.

Chaparro, A., Young, R.L.S. (1993). Reading With Rods: The Superiority of Central Vision for Rapid Reading *Investigative Ophthalmology & Visual Science,* **34,** 2341-2347.

Chastain, G. (1983). Task and contrast effects on perform-ance with parafoveal stimulus pairs. *Psychological Re-search,* **45,** 147-156.

Chastain, G. (1986). Evidence for feature perturbations from character misidentification. *Perception & Psychophysics,* **39,** 301–306.

Chun, M. M., Wolfe, J.M. (1996). Just Say No: How Are Visual Searches Terminated When There Is No Target Present? *Cognitive Psychology,* **30,** 39-78.

Chung, S. T. L., Levi, D. M. and Legge, G. E. (2001). Spatial-frequency and contrast proprieties of crowding *Vision Research,* **41,** 1833-1850.

Chung, S. T. L., Mansfield, J. S., & Legge, G. E. (1998). Psychophysics of reading. XVIII. The effect of print size on reading speed in normal peripheral vision. *Vision Research,* **38,** 2949–2962.

Cowey, A. and Rolls, E. T. (1974). Human cortical magnification factor and its relation to visual acuity *Expl. Brain Res.,* **21** 447-454

Cytowic, R. E. (1989). Synaesthesia: a Union of the Senses.

Cytowic, R. E. (1993) *The Man Who Tasted Shapes.*

Cytowic, R. E. (1996). Synaesthesia: Phenomenology and Neuropsychology. A review of Current Knowledge. *Psyche,* **2**.

Cytowic, R. E. (1997). Synaesthesia: Phenomenology and neuropsychology. In S. Baron-Cohen, & J. E.

Harrison (Eds.), Synaesthesia: Classic and contemporary readings**,** 17–39.

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

Daniel, P. M. and Whitteridge, D. (1961). The representation of the visual field on the cerebral cortex in monkeys *Journal of Physiology, Lond.,* **159** 203-221

Dixon, M. J., Smilek, D., Cudahy, C.& Merikle, P.M. (2000). Five plus two equal yellow. *Nature,* **406,** 365.

Dosher, B. A. (1976). The retrieval of sentences from memory: A speed-accuracy study. *Cognitive Psychology.,* **8,** 291-310.

Dosher, B. A. (1979). Empirical approaches to information processing: Speed-accuracy tradeoff functions or reaction time: A reply. *Acta Psychologica.,* **43,** 347-359.

Dosher, B. A. R., G. (1991). Judgments of semantic and episodic relatedness: Common time-course and failure of segregation. *Journal of Memory and Language.,* **30,** 125-160.

Duncan, J. and Humphreys, G. W. (1989). Visual search and stimulus similarity *Psychological Review,* **96,** 433-458.

D'Zmura, M. (1991). Color in visual search *Vision Research,* **31,** 951-966.

Edquist, J., Rich, A.N., Brinkman, C., Mattingley, J.B. (2006). Do synaesthetic colours act as unique features in visual search? *Cortex,* **42,** 222-31.

Efron, B., Tibshirani, R.J. (1993) *An introduction to the bootstrap.*, New York.

Ehlers, H. (1936). The movements of the eyes during reading. *Acta Ophthalmologica,* **14,** 56-63.

Estes, W. K. (1972). Interaction of signal and background variables in visual processing. *Perception & Psychophysics,* **12,** 278–286.

Estes, W. K. (1974). Redundancy of noise elements and signals in visual detection of letters. *Perception & Psychophysics,* **16,** 53–60.

Felisberti, F., Solomon, J. A. and Morgan, M. M. (2005). Crowding and salience *Journal of Vision*.

Fitzpatrick, D. (2000). Seeing beyond the receptive field in primary visual cortex. *Current Opinion in Neurobiology.,* **10,** 438-443.

Flom, M. C., Heat, G. and Takahashi, E. (1963). Contour interaction and visual resolution: Contralateral effect *Science,* **142,** 979-980.

Foley, J. M. (1994). Human luminance pattern-vision mechanisms: masking experiments require a new model. *J Opt Soc Am,* **11,** 1710-9.

Foster, D. H. and Ward, P. A. (1991b). Horizontal-vertical filters in early vision predict anomalous line-orientation identification frequencies. *Proc. R. Soc. Lond.,* **243,** 83-86.

Foster, D. H., Ward, P. A. (1991a). Asymmetries in oriented-line detection indicate two orthogonal filters in early vision. *Proceedings of the Royal Society (London B),* **243,** 75-81.

Galton, F. (1880). Visualised numerals *Nature,* **21,** 252-256.

Geiger, G., Lettvin, J.Y. (1986). Enhancing the perception of form in peripheral vision. *Perception,* **15,** 119-30.

Gheri, C., Chopping, S., Morgan, M.J. (2008). Synaesthetic colours do not camouflage form in visual search. *Proc Biol Sci,* **275,** 841-6.

Gibson, J. J. and Radner, M. (1937). Adaptation and contrast in the perception of tilted lines. *Journal of Experimental Psychology,* **20,** 453-469.

Gilbert, C. D. (1998). Adult cortical dynamics. *Physiological Reviews.,* **78,** 467-485.

Green, B. F., Anderson, L. K. (1956). Color coding in a visual search task. *J. Exp. Psychol.,,* **51,** 19-24.

Green, D. M. and Swets, J. A. (1966) *Signal Detection Theory and Psychophysics,* Wiley, New York.

Grossenbacher, P. G. and Lovelace, C. T. (2001). Mechanisms of synesthesia: cognitive and physiological constraints. *Trends Cogn Sci,* **5,** 36-41.

He, S., Cavanagh, P. and Intriligator, J. (1996). Attentional resolution and the locus of visual awareness *Nature,* **383,** 334-337.

Heeger, D. J. (1992). Half-squaring in responses of cat striate cells *Visual Neuroscience,* **9,** 427-423.

Herzog, M. H., Koch, C. (2001). Seeing properties of an invisible object: Feature inheritance and shine-through. *Proc Natl Acad Sci U S A,* **98,** 4271–4275.

Hubbard, E. M., & Ramachandran, V. S. (2005a). Neurocognitive mechanisms of synesthesia. *Neuron,* **48,** 509-520.

Hubbard, E. M., Arman, A. C., Ramachandran, V. S., & Boynton, G. M. (2005b). Individual differences among grapheme-color synesthetes: brain-behaviour correlations. *Neuron,* **45,** 975-985.

Huckauf, A., & Heller, D. (2002). What various kinds of errors tell us about lateral masking effects. *Visual Cognition,* **9,** 889-910.

Intriligator, J. and Canavagh, P. (2001). The spatial resolution of Visual attention *Cognitive Psychology,* **43,** 171-216.

Kapadia, M. K., Ito, M., Gilbert, C. D., & Westheimer, G. (1995). Improvement in visual sensitivity by changes in local context: Parallel studies in human observers and in V1 of alert monkeys. *Neuron,* **15,** 843-856.

Kim, C. Y., Blake, R., & Palmeri, T. (2006). Perceptual interaction between real and synestehtic colors. *Cortex*.

Kinchla, R. A. (1974). Detecting target elements in multielement arrays: A comfusability model. *Perception & Psychophysics,* **15,** 149-158.

Kooi, F. L., Toet, A., Tripathy, S. P. and Levi, D. M. (1994). The effect of similarity and duration on spatial interaction in peripheral vision *Spatial Vision,* **9,** 255-279.

Kooi, F. L., Toet, A., Tripathy, S. P., & Levi, D. M. (1994). The effect of similarity and duration on spatial interaction in peripheral vision. *Spatial Vision,* **8,** 255-279.

Korte, W. (1923). Über die Gestaltauffassung imindirekten Sehen *Zeitschrift für Psychologie,* **93,** 17-82.

Laeng, B., Svartdal, F., and Oelmann, H. (2004). Does colour synaesthesia pose a paradox for early-selection theories of attention? *Psychol. Sci.,* **15,** 277-281.

Latham, K., Whitaker, D. (1996a). A comparison of word recognition and reading performance in foveal and peripheral vision. *Vision Research,* **36,** 2665-74.

Legge, G. E., Ahn, S.J., Klitz, T.S., Luebeker A. (1997). Psychophysics of Reading--XVI. The Visual Span in Normal and Low Vision *Vision Research,* **37,** 1999-2010.

Legge, G. E. and Foley, J. M. (1980). Contrast making in human vision *Journal of the Optical Society of America,* **70,** 1458-1471.

Legge, G. E., Mansfield, J. S. and Chung, S. T. L. (2001). Psychophysics of reading XX. Linking letter recognition to reading speed in central and peripheral vision. *Vision Research,* **41,** 725-743.

Legge, G. E., Rubin, G. S., Pelli, D. G. and Schleske, M. M. (1983). Psychophysics of reading: II. Low vision *Vision Research,* **25,** 253-266.

Legge, G. E., Rubin, G. S., Pelli, D. G., & Schleske, M. M. (1985). Psychophysics of reading. II. Low vision. *Vision Research,* **25,** 253–266.

Levi, D., Klein, S. and Aitsabaomo, A. P. (1985). Vernier acuity, crowding and cortical magnification factor *Vision Research,* **25,** 963-967.

Levi, D. M., Hariharan, S. and Klein, S. A. (2002a). Suppressiveand facilitatory spatial interactions in peripheral vision: Peripheral crowding is neither size invariant nor simple contrast masking. *Journal of Vision,* **2,** 167-177.

Levi, D. M. and Klein, S. A. (1985). Vernier acuity, crowding and amblyopia. *Vision Research,* **25,** 979-991.

Levi, D. M., Klein, S. A. and Hariharan, S. (2002b). Suppressive and facilitation spatial interaction in foveal vision: Foveal crowding is simple contrast masking. *Journal of Vision,* **2,** 140-166.

Levi, D. M., Song, S., Pelli, D.G. (2007). Amblyopic reading is crowded. *J Vis,* **7(2),** 1-17.

Li, W., Thier, P. and Wehrhahn, C. (2000). Contextual influence on orientation discrimination of humans and responses of neurons in V1 of alert monkeys *J Neurophysiol,* **83,** 941-54.

Li, Z. (2002). A saliency map in primary visual cortex *Trends in Cog Sci,* **6,** 9-16.

Liu, L., & Arditi, A. (2000). Apparent string shortening concomitant with letter crowding. *Vision Research,* **40,** 1059-1067.

Livne, T. and Sagi, D. (2007). Configuration influence on crowding *Journal of Vision,* **7(2),** 1-12.

Lund, J. S., Yoshioka, T., & Levitt, J. B. (1993). Comparison of intrinsic connectivity in different areas of macaque monkey cerebral cortex. *Cortex,* **3,** 148-162.

Luria, A. R. (1968). The mind of a mnemonist.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin,* **109,** 163-203.

MacLeod, C. M., Dunbar, K. (1988). Training and Stroop-like interference: Evidence for a continuum ofautomaticity. *Journal of Experimenta Psychology: Learning, Memory, and Cognition,* **10,** 304-315.

Malania, M., Herzog, M., & Westheimer, G. (2007). Grouping of contextual elements that affect vernier thresholds *Journal of Vision,* **7,** 1–7.

Mansfield, J. S., Legge, G. E., & Ortiz, A. (1998). The role of segmentation in lateral masking. *Investigative Ophthalmology & Visual Science,* **39,** 859.

Mark, L. E. (1975). On colored-hearing synesthesia: cross-modal translations of sensory dimensions. *Psyhol Bull,* **82,** 303-331.

Marks, L. E. (2000). in Varieties of Anomalous Experience: Examining the Scientific Evidence. (eds Cardena, E., Lynn, S.J. & Krippner, S. C.)**,** 121-149.

Mattingley, J. B., Rich, A. N., Yelland, G. and Bradshaw, J. L. (2001). Unconscious priming eliminates automatic binding of colour and alphanumeric form in synaesthesia. *Nature,* **410,** 580-582.

McElree, B., Carrasco, M. (1999). The Temporal Dynamics of Visual Search: Evidence for Parallel Processing in Feature and Conjunction Searches *Journal of Esperimental psychology: Human Perception and Performance,* **25,** 1517-1539.

Mills, C. B. and el., e. (1999). Digit synaesthesia: a case study using a Stroop-type test. *Cognit Neuropsychol,* **16,** 181-191.

Morgan, M. J., Adam, A. and Mollon, J. D. (1992). Dichromats break colour-camouflage of textural boundaries *Proc. Roy. Soc.,* **B. 248,** 291-295.

Morgan, M. J., Castet, E. and Ward, R. M. (1998). Visual search for a tilted target: tests of the spatial uncertainty model *Quarterly J. Exp. Psychol.,* **51,** 347-371.

Nagy, A. L. and Sanchez, R. R. (1990). Critical colour differences determined with visual search task. *J Opt Soc Am A,* **7,** 1209-1217.

Nagy, A. L., Sanchez, R. R. and Hughes, T. C. (1990). Visual search for color differences with foveal and peripheral vision. *J. Opt. Soc. Am. - A,* **7,** 1995-2001.

Nakayama, K. and Silverman, G. H. (1986a). Serial and parallel processing of visual feature conjunctions *Nature,* **320,** 264-265.

Nakayama, K. a. G. H. S. (1986b). Serial and parallel encoding of feature conjunction *Invest. Ophthal. and Visual Science,* **27,** 182 (Supplement).

Nazir, T. A. (1992). Effects of lateral masking and spatial precueing on gap-resolution in central and peripheral vision. *Vision Research,* **32,** 771-777.

Neisser, U. (1963). Decision time without reaction time: experiments in visual scanning. *American Journal of Psychology,* **76,** 376-385.

Neisser, U. (1967) *Cognitive Psychology.*

Nothdurft, H. C. (2000a). Salience from feature contrast: variations with texture density *Vision Research,* **40,** 3181-3200.

Nothdurft, H. C. (2000b). Salience from feature contrast: temporal properties of saliency mechanisms *Vision Research,* **40,** 2421-2435.

Nothdurft, H. C. (2002). Attention shifts to salient targets *Vision Research,* **42,** 1287-1306.

Nunn, J. A., Gregory, L. J., Brammer, M., Williams, S. C., Parslow, D. M., Morgan, M. J., Morris, R. G., Bullmore, E. T., Baron-Cohen, S. and Gray, J. A. (2002). Functional magnetic resonance imaging of synesthesia: activation of V4/V8 by spoken words *Nat Neurosci,* **5,** 371-5.

Odgaard, E. C., Flowers, J. H., Bradman, H. L. (1999). An investigation of the cognitive and perceptual dynamics of a colour-digit synaesthete. *Perception,* **28,** 651-664.

Over, R., Broerse, J. and B.Crassini (1972). Orientation Illusion and masking in central and peripheral vision. *Journal of Experimental Psychology,* **96,** 25-31.

Palmer, J. (1994). Set-Size effects in visual search: the effect of attention is independent of the stimulus for simple tasks *Vision Research,* **32,** 1703-1721.

Palmer, J., Ames, C. T. and Lindsey, D. T. (1993). Measuring the effect of attention on simple visual search *JEP:HPP,* **19,** 108-130.

Palmer, J., Verghese, P. and Pavel, M. (2000). The psychophysics of visual search *Vision Res,* **40,** 1227-68.

Palmeri, T., Blake, R., Marois, R., Flanery, M., & Whetsell, W. (2002). The perceptual reality of synaesthetic colors. *Proc Natl Acad Sci U S A,* **99**.

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

Paulesu, E. (1995). The physiology of coloured hearing: a PET activation study of colour-word Synasthesia. *Brain,* **118,** 661-667.

Pelli, D. G. (1985). Uncertainty explains many aspects of visual contrast detection and discrimination *Journal of the Optical Society of America A,* **2,** 1508-1532.

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

Pelli, D. G., & Farell, B. (1999). Why use noise? *Journal of the Optical Society of America A,* **16,** 647-653.

Pelli, D. G., Palomares, M. and Majaj, N. J. (2003). Crowding is unlike ordinary masking: distinguishing feature edtection and integration *Journal of Vision*.

Pelli, D. G., Palomares, M. and Majaj, N. J. (2004). Crowding is unlike ordinary masking: distinguishing feature edtection and integration *Journal of Vision*.

Pelli, D. G., Palomares, M., Majaj, N.J. (2004). Crowding is unlike ordinary masking: distinguishing feature integration from detection *J Vis,* **4,** 1136-69.

Pelli, D. G., Tillman, K.A., Freeman, J, Su, M., Berger, T.D., Majaj, N.J. (2007). Crowding and eccentricity determine reading rate. *J Vis,* **7(2),** 1-36.

Petrov, Y., & Popple, A. V. (2007). Crowding is directed to the fovea and preserves only feature contrast. *Journal of Vision,* **7,** 1-9.

Poder, E. (2006). Crowding, feature integration and two kind of 'attention' *Journal of Vision,* **6,** 163–169.

Polat, U., & Sagi, D. (1993). Lateral interactions between spatial channels: Suppression and facilitation revealed by lateral masking experiments. *Vision Research,* **33,** 993-999.

Polat, U. and Sagi, D. (1994). The architecture of perceptual spatial interactions *Vision Research,* **34,** 73-78.

Potter, M. C. (1984) In *New methods in reading comprehension research*Hillsdale, pp. 91-118.

Ramachandran, V., & Hubbard, E. M. (2001a). Synaesthesia: a window into perception, thought and language. *J. consciousness Studies,* **8,** 3-34.

Ramachandran, V., & Hubbard, E. M. (2003). The phenomenology of Synaesthesia. *Journal of Consciousness Studies,* **10,** 49-57.

Ramachandran, V. S. and Hubbard, E. M. (2001b). Psychophysical investigations into the neural basis of synaesthesia *Proc R Soc Lond B Biol Sci,* **268,** 979-83.

Rich, A., & Mattingley, J. (2002). Anomalous perception in synaesthesia: a cognitive neuroscience perspective. *Nat Rev Neurosci.,* **3,** 43-52.

Rich, A. N. and Mattingley, J. B. (2002). Anomalous perception in synaesthesia: A cognitive neuroscience perspective. *Nature Reviews Neuroscience,* **3,** 43–52.

Rockland, K. S., & Lund, J. S. (1983). Intrinsic laminar lattice connections in primate visual cortex. *Journal of Comparative Neurology,* **216,** 303-318.

Ross, J., Speed, H.D., Morgan, M.J. (1993). The effects of adaptation and masking on incremental thresholds for contrast. *Vis Res,* **33,** 2051-6.

Sagiv, N., Heer, J., Robertson, L. (2006). Does binding of synesthetic color to the evoking grapheme require attention? *Cortex,* **42,** 232-42.

Schiltz, K., Trocha, K., Wieringa, B.M., Emrich, H.M., Johannes, S., Münte, T.F. (1999). Neurophysiological aspects of synesthetic experience. *J Neuropsychiatry Clin Neurosci.,* **11,** 58-65.

Schneider, W., Shiffrin, R.M. (1977). Controlled and Automatic Human Information Processing: 1. Detection, Search, and Attention *Psychological Review,* **84,** 1-66.

Seron, X., Pesenti, M. and Noel, M. P. (1992). Images of numbers, or 'When 98 is upper left and 6 sky blue'. *Cognition,* **44,** 159-196.

Shani, R., Sagi, D. (2005). Eccentricity effects on lateral interactions. *2005,* **45,** 2009-24.

Shaw, P. (1969). Processing of tachistoscopic displays with controlled order of characters and spaces. *Perception & Psychophysics,* **6,** 257-266.

Shiffrin, R. M., Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review,* **84,** 127-190.

Smilek, D., Dixon,M.J.,Cudahy,C.&Merikle,P.M. (2001). Synaesthetic photisms influence visual perception *J. Cogn. Neurosci.,* **13,** 930-936.

Snowden, R. J., Hammett, S.T. (1998). The effects of surround contrast on contrast thresholds, perceived contrast and contrast discrimination *Vision Research,* **38,** 1935–1945.

Solomon, J. A., Felisberti, F. M. and Morgan, M. J. (2004). Crowding and tilt illusion: towards a unified account. *Journal of Vision*.

Solomon, J. A. and Morgan, M. J. (2001). Odd-men-out are poorly localized in brief displays *Investigative Ophthalmology & Visual Science,* **42,** 4973.

Solomon, J. A., Morgan MJ (2006). Stochastic re-calibration: contextual effects on perceived tilt. *Proc Biol Sci.,* **22,** 1601.

Solomon, J. A., Watson, A. B. and Morgan, M. J. (1999). Transducer model produces facilitation from opposite-sign flanks *Vision Res,* **39,** 987-92.

Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. *Acta Psychologica,* **30**.

Strasburger, H. (2005). Unfocussed spatial attention underlies the crowding effect in indirect form vision. *Journal of Vision,* **5(11),** 1024-2037.

Strasburger, H., Harvey, L. O. and Rentschler, I. (1991). Contrast thresholds for identification of numeric characters in direct and eccentric view. *Percept Psychophys,* **49,** 495-508.

Stuart, J. A. and Burian, H. M. (1962). A study of separation difficulty:its relationship to visual acuity in normal and amblyopic eyes. *American Journal of Ophthalmology,* **53,** 471.

Takahashi, E. (1967). Effect of flanking contours on visual resolution at foveal and near-foveal loci.

Taylor, S. G., & Brown, D. R. , (1972). Lateral visual mask-ing: Supraretinal effects when viewing linear arrays with unlimited viewing time. *Perception &*

Psychophysics, **12,** 97-99.

Toet, A. and Levi, D. M. (1992). The two-dimensional shape of spatial interaction zones in the parafovea *Vision Research,* **32,** 1349-1357.

Townsend, J. T., Taylor, S. G. and Brown, D. R. (1971). Lateral masking for letters with unlimited viewing time. *Perception and Psychophysics,* **10,** 375-378.

Treisman, A. (1985). Preattentive processing in vision *Computer Vision, Graphics, and Image Processing,* **31,** 156-177.

Treisman, A. and Gormican, S. (1988). Feature analysis in early vision: evidence from search asymmetries *Psychological Review,* **95,** 15-48.

Treisman, A. and Schmidt, H. (1982). Illusory conjunctions in the perception of objects. *Cognitive Psychology,* **14,** 107–141.

Treisman, A. and Souther, J. (1985). Search asymmetry: a diagnostic for preattentive processing of separable features *Journal of Experimental Psychology, General,* **114,** 285-310.

Treisman, A. M. and Gelade, G. (1980). A Feature-integration theory of attention *Cognitive Psychology,* **12,** 97-136.

Treisman, A. M. and Sato, S. (1990). Conjunction search revisited *Journal of Experimental Psychology, Human Perception and Performance,* **16,** 459-478.

Tripathy, P. S. and Cavanagh, P. (2002). The extent of crowding in peripheral vision does not scale with target size *Vision Research,* **42,** 2357-2369.

Van Essen, D. C., Newsome, W. T. and Maunsell, J. H. R. (1984). The visual field representation in striate cortex of the macaque monkey: asymmetries, anisotropies, and individual variability *Vision Research,* **24,** 429-448.

Virsu, V. and Rovamo, J. (1979). Visual resolution, contrast sensitivity, and the cortical magnification factor *Experimental Brain Research,* **37,** 475-494.

Ward J, L. R., Salih S, Sagiv N. (2007). Varieties of grapheme-colour synaesthesia: a new theory of phenomenological and behavioural differences. *Conscious Cogn.,* **16,** 913-31.

Ward, J., Simner, J. (1993). Lexical-gustatory synaesthesia: linguistic and conceptual factors *Cognition,* **89,** 237-261.

Watson, A. B. and Pelli, D. G. (1983). QUEST: A Bayesian adaptive psychometric method *Perception and Psychophysics,* **33,** 113-120.

Westheimer, G. and Hauske, G. (1975). Temporal and spatial interference with vernier acuity *Vision Res,* **15,** 1137-41.

Westheimer, G., Shimamura, K. and McKee, S. P. (1976). Interference with line-orientation sensitivity *J Opt Soc Am,* **66,** 332-8.

Wilken, P. and Ma, W. J. (2004). A detection theory account of change detection *Journal of Vision,* **4,** 1120-1135.

Wilkinson, F., Wilson, H. R. and Ellemberg, D. (1997). Lateral interactions in peripherally viewed texture arrays *J Opt Soc Am A,* **14,** 2057-68.

Williams, C. B. and Hess, R. F. (1998). Relationship between facilitation at threshold and suprathreshold contour integration. *Journal of the Optical Society of America A, Optics, Image Science, and Vision,,* **15,** 2046–2051.

Wilson, H. R., Wilkinson, F. and Asaad, W. (1997). Concentric orientation summation in human form vision *Vision Res,* **37,** 2325-30.

Wolfe, J. M. (1998) *Visual Search*.

Wolfe, J. M., Friedman-Hill, S. R. and Bilsky, A. B. (1994). Parallel Processing of part/whole information in visual search task. *Percept Psychophys,* **55,** 537-550.

Wolford, G., & Chambers, L. (1983). Lateral masking as a function of spacing. *Perception & Psychophysics,* **33,** 129-138.

Wolford, G. and Chambers, L. (1984). Contour interaction as a function of retinal eccentricity *Percept Psychophys,* **36,** 457-460.

Wollen, K. A., & Ruggiero, F. T. (1983). Coloured-letter synaesthesia. *Journal of Mental Imagery,* **7,** 83-86.

Wollen, K. A. and Ruggiero, F. T. (1983). Colored-letter synesthesia. *J Mental Imagery,* **7,** 83-86.

Woodrow, H. (1938). The effect of pattern upon simulta-neous letter-span. *American Journal of Psychology,* **51,** 83-96.

Yeshurun, Y. and Carrasco, M. (1998). Attention improves or impairs visual performance by enhancing spatial resolution *Nature,* **396,** 72-5.

Yeshurun, Y. and Carrasco, M. (1999). Spatial attention improves performance in spatial resolution tasks *Vision Res,* **39,** 293-306.

Yeshurun, Y. and Carrasco, M. (2000). The locus of attentional effects in texture segmentation *Nat Neurosci,* **3,** 622-7.

Yu, C., Klein, S.A., Levi, D.M. (2002). Facilitation of contrast detection by cross-oriented surround stimuli and its psychophysical mechanisms. *J Vis,* **2,** 243-55.