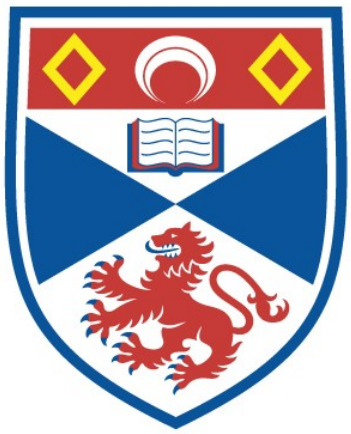
PSYCHOPHYSICAL STUDIES OF INTERACTIONS BETWEEN LUMINANCE AND CHROMATIC INFORMATION IN HUMAN VISION

Stéphane Clery

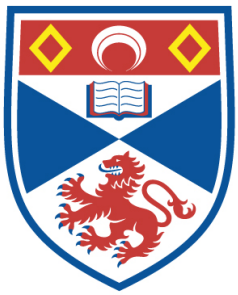
A Thesis Submitted for the Degree of PhD
at the
University of St Andrews



2014

Psychophysical studies of interactions between luminance and chromatic information in human vision

Stéphane Clery



University of
St Andrews

This thesis is submitted to the University of St Andrews for the degree of PhD

28th October 2013

1. Candidate's declarations:

I, Stéphane Clery, hereby certify that this thesis, which is approximately 75 000 words in length, has been written by me, that it is the record of work carried out by me and that it has not been submitted in any previous application for a higher degree.

I was admitted as a research student in September 2009 and as a candidate for the degree of Doctor of Philosophy in September 2009; the higher study for which this is a record was carried out in the University of St Andrews between 2009 and 2013.

Date 25/10/2013 signature of candidate _____

2. Supervisor's declaration:

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of Doctor of Philosophy in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

Date 25/10/2013 signature of supervisor _____

3. Permission for electronic publication: (*to be signed by both candidate and supervisor*)

In submitting this thesis to the University of St Andrews I understand that I am giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. I also understand that the title and the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that my thesis will be electronically accessible for personal or research use unless exempt by award of an embargo as requested below, and that the library has the right to migrate my thesis into new electronic forms as required to ensure continued access to the thesis. I have obtained any third-party copyright permissions that may be required in order to allow such access and migration, or have requested the appropriate embargo below.

The following is an agreed request by candidate and supervisor regarding the electronic publication of this thesis:

(iii) Embargo on both all of printed copy and electronic copy for the same fixed period of two years on the following ground(s):

publication would preclude future publication.

Date 25/10/2013 signature of candidate _____ signature of supervisor _____

A supporting statement for a request for an embargo must be included with the submission of the draft copy of the thesis. Where part of a thesis is to be embargoed, please specify the part and the reasons.

Abstract

In this thesis, I investigated how human vision processes colour and luminance information to enable perception of our environment. I first tested how colour can alter the perception of depth from shading. A luminance variation can be interpreted as either variation of reflectance (patterning) or variation of shape. The process of shape-from-shading interprets luminance variation as changes in the shape of the object (e.g. the shading on an object might elicit the perception of curvature). The addition of colour variation is known to modify this shape-from-shading processing. In the experiments presented here I tested how luminance driven percepts can be modified by colour. My first series of experiments confirmed that depth is modulated by colour. I explored a larger number of participants than previously tested. Contrary to previous studies, a wide repertoire of behaviour was found; participants experienced variously more depth, or less depth, or no difference. I hypothesised that the colour modulation effect might be due to a low-level contrast modulation of luminance by colour, rather than a higher-level depth effect. In a second series of experiments, I therefore tested how the perceived contrast of a luminance target can be affected by the presence of an orthogonal mask. I found that colour had a range of effects on the perception of luminance, again dependant on the participants. Luminance also had a similar wide range of effects on the perceived contrast of luminance targets. This showed that, at supra-threshold levels, a luminance targets contrast can be modulated by a component of another orientation (colour or luminance defined). The effects of luminance and colour were not following a particular rule. In a third series of experiments, I explored this interaction at detection levels of contrast. I showed cross-interaction between luminance target and mask but no effects of a colour mask.

Acknowledgements

Firstly, I would like to thank my supervisor, Julie Harris, from whom I have learned greatly. Thanks for being always available when I had "a quick question". I would also like to thank Marina Bloj for the helpful support and interesting discussions. Thanks to George and Glenn for being outstanding co-workers. Thanks to all the vision lab members past and present for insightful discussions during journal club and good camaraderie outside the lab. Thanks to Anastasia Kolokolnikova, for your assistance on data collection for the colour-shading experiment. Thanks to the EPSRC for funding our research. I would also like to thank my former professors, teachers and supervisors who helped and inspired me along the way. I would like to thank all the members of staff at the School of Psychology & Neuroscience for always being friendly and professional. Thanks also to the academic staff for creating an inspiring and stimulating environment. I am indebted to the staff and doctors at Ninewells hospital, thank you for your work. Finally, I would like to thank my friends, and family for being supportive during the most difficult parts and for believing in me, *mille mercis*. Thank you Jessica for your unflinching support these past years.

Contents

List of Figures

List of Tables

List of abbreviations

2 IFC: two-interval forced choice
CIE: Commission International de l´Éclairage
CRF: classical receptive field
CSF: contrast sensitivity function
DOG: difference of Gaussian model
H-K effect: Helmholtz–Kohlrausch effect
ERG: Electroretinogram
LGN: lateral geniculate nucleus
M (cells/layers): Magnocellular (cells/layers)
P (cells/layers): Parvocellular (cells/layers)
K (cells/layers): Koniocellular (cells/layers)
RF: receptive field
SF: spatial frequency
SFS: shape-from-shading
TvC: threshold versus (mask) Contrast
VS: visual system
XOE: cross-orientation enhancement
XOI: cross-orientation interaction
XOM: cross-orientation masking

Glossary

Visual channel:

A visual channel is a perceptual sub-organization of visual information. There are three main pathways of visual information (see Chapter 3). Their operational definition consists of a recombination of retinal signals (L, M and S) into three distinct signals M-L, (M+L)-S & L+M, perceptually equivalent to red-green, yellow-blue, and luminance perception. As such, these channels are also called post-receptoral mechanisms. See Chapter 4 for an extended definition of photometry and colorimetry essential for the characterization of luminance and chromatic channels.

Visual pathway:

A visual pathway is the physiological sub-organization that carries visual information in the brain. It carries but also modifies the information by performing additional computations. There are three main visual pathways starting from the retina (P, M and K, see Chapter 2 for details). Although it might have been suggested in the past, there is no 1 to 1 relationship between visual channels (psychophysically defined, see previous definition) and the main visual pathways. This fact is essential to understand the interactions between channels and signals presented in this thesis.

Chapter 1

General introduction

The processing of visual information by the human visual system involves a dedicated part of the central nervous system from the retina through the lateral geniculate nucleus (LGN) to Visual area 1 (V1) and higher cortical areas. This visual processing involves physiological pathways carrying specific types of information whose properties can be traced back to the ganglion cells in the retina. There are 3 physiological pathways of information processing carrying visual input to the cortex. A more detailed and in depth introduction to the physiology of colour and luminance vision (in the context of spatial information) will be given in Chapter 2.

Before the physiological details of visual processing were known, psychophysical channels of visual information had been suggested (e.g. M. D. Wright, 1946). A common suggestion is that three fundamental psychophysical channels exist; one channel carrying information about luminance and two about chromatic information. A psychophysical channel is defined operationally, through the use of psychophysical experiments. The standardized psychophysical metrics used in this thesis will be reviewed in Chapter 4 (General Methods). It was found that luminance and chromatic channels had different properties in the way they respond to spatial and temporal information, for example they process visual information at different scales (e.g. Mullen, 1985). Channels can be said to be composed of orientation and spatial scale sub-channels. These concepts and data relating to them will be presented in the psychophysical review chapter (Chapter 3). A particular subject of interest is how these channels interact under differing conditions. This is studied because visual processing plays an important role for us to interpret and interact with our environment. One of the questions of this thesis is about the functional role of these channels and interactions between them. In Chapters 2 and 3, I will review how orientation processing (part of the processing of edges) is achieved by interacting channels. In the experimental chapters, I will explore how luminance and colour interact in the processing of information at different levels. Furthermore, I will contrast results when possible with results from the processing of luminance-luminance interactions.

In the first experimental chapter (Chapter 5), I will study the 'colour-shading effect' (Kingdom, 2003), a phenomenon that involves modulation of achromatic shape-from-

shading (SFS) by isoluminant chromatic information. I will explore this phenomenon in a larger number of participants than previously tested. In the second experiment (Chapter 6), I will present data on how orthogonal components can modulate the perception of suprathreshold contrast. The effect of mask types (sinusoids with smooth edges or square-wave with straight edges) will be studied in a population of 12 participants. These orthogonal masks were defined both in luminance and in red-green isoluminant colours. The effects of target contrast, mask contrast and mask type (luminance or chromatic) will be detailed for 22 participants. In a third experiment (Chapter 7) the effects of orthogonal modulation upon the detection of a luminance target will be detailed. The masks could either be luminance or colour defined in this study.

In order to define stimuli that are isoluminant for each participant, experiments are required. The nominal standards for luminance are not sufficient for this type of study (see Chapter 4). Consequently my participants did perform isoluminance experiments. Different methods were used and as I used some variations in my stimuli the data obtained have some interesting characteristics in terms of methodology, this data is presented in a dedicated chapter (Chapter 8). The results of this thesis are discussed in the final discussion chapter (Chapter 9).

Chapter 2

Physiology of luminance and chromatic processing from the retina to the cortex

2.1 Introduction

The aim of this introductory chapter is to present a short description of the physiological basis of information processing through separate pathways. This will be done with an emphasis on luminance and colour processing. In this chapter, I explore perception in terms of signal processing, starting from the retinae to the visual cortex, thereby focusing on the physiological mechanisms of perception. I will detail current knowledge of how luminance, chromatic, and spatial information are computed from the retina to the cortex. The following chapter, Chapter 3, follows the natural consequence of visual processing which is conscious perception, and is therefore oriented towards psychophysics. Chapter 3 will also focus on image/scene processing.

2.2 Retina

In cross-section, the retina is a layered structure, with the photoreceptor cells outer segment at the most external layer. Other layers contain receptor bodies, horizontal cells, ganglion cells, and optic nerve fibres (for a more complete description see classic work by Polyak, 1941, 1957; Ramon y Cajal & Felipe, 1892). Visual processing starts when light[1] reaches the back of the eye, goes through the retinal layers to eventually be absorbed by photoreceptors; which then gets processed by intermediate cells followed by ganglions projecting to the cortex via the optic nerve.

I will first describe details of the cone photoreceptors, then the retinal processing done by intermediate and ganglion cells.

[1] In the General Methods chapter, Chapter 4, I will explore in more detail the properties of light emission and how it is measured.

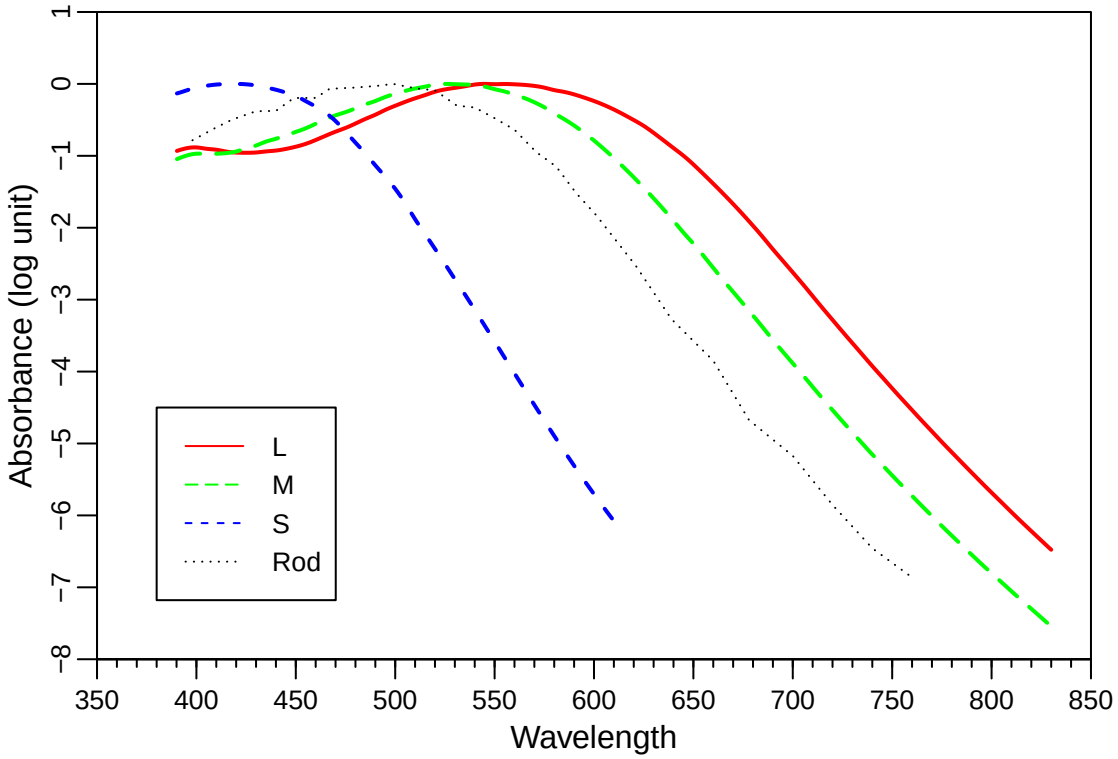


Figure 2.1: Absorbance spectra of cones and rods in humans, estimates by Stockman and Sharpe (2000, cones) and Kraft et al. (1993, rods).

2.2.1 Cones

Primates' eyes have evolved two photoreceptors types: the cones and the rods. The cones are responsible for daylight viewing (photopic condition) and the rods are responsible for low-light vision (scotopic condition, see Chapter 4, General methods, for more details).

Humans and some other primates possess three types of cones, and are, therefore, referred to as trichromats. The three types of cones are called L, M, and S;[2] for long, medium, and short absorption range (see Figure 2.1). Having multiple cone types means that at the initial processing step there are multiple (three) concurrent representations of the visual input (Lennie, 2003). The cone mosaic varies with eccentricity; cone density is highest at the fovea. In terms of sampling, trichromats have on average, four times more L cones than M cones and even fewer S cones (Kremers et al., 2000; Roorda, Metha, Lennie, & Williams, 2001). The approximate ratios are 20:5:1 for L:M:S (Roorda et al., 2001).

As we will see further on, each of the cone types are not colour detectors *per se*. L (red) is not used as a photo-detector for red. As can be seen by their name, their main difference stands in the peak absorption wavelength (Stockman & Sharpe, 2000).

In the rest of this chapter, I will focus on cone processing as this thesis is mainly

[2]The three cone types used to be referred to as Red, Green and Blue; however, these terms are now obsolete and can be misleading.

interested in processing under photopic conditions.

Transduction and univariance Photoreceptors absorb photons stochastically; absorption probability varies depending on the opsin (photo-reactive protein) and the wavelength of the light. Figure 2.1 shows the absorbance probability of humans' four receptor types. Once a photon is absorbed this information is *transduced*[3] into voltage signals in the receptors' membrane (Baylor, 1987; McNaughton, 1990; Yau, 1994). Photoreceptors are non-spiking neurons and release the neurotransmitter glutamate (Brandon & Lam, 1983; Copenhagen & Jahr, 1989), that affects the subsequent firing rates of retinal neurons.

Concretely, once a photon is absorbed a photoreceptor membrane will hyperpolarize (Baylor, Hodgkin, & Lam, 1974). Importantly as stated by Naka and Rushton (1966, p. 538), "each visual pigment can only signal the rate at which it is effectively catching quanta; it cannot also signal the wave-length associated with the quanta caught". This is called *the principle of univariance*. Once a photon (or series of photons) is (/are) absorbed, the output signal of a single photoreceptor cannot be decoded to work out either the wavelength or intensity (flux) of light; the two sources of information are combined. In the normal range, the response of the cone is linearly dependant on the intensity of the light (Schnapf, Nunn, Meister, & Baylor, 1990), or number of photons absorbed. At higher values, the response saturates. Consequently, the output current waveform of the photoreceptor is preserved and scaled with light intensity.

By reading an L cone output (see Figure 2.1) we cannot distinguish between an intense (monochrome) blue light and a less intense red (monochrome) light as the output only depends "on the number of photons absorbed and not the wavelength of the absorbed photons" (Kraft et al., 1993, p. 754). As cone output is agnostic to wavelength, the experimenter can manipulate light wavelength and intensity to match a specific cone response shape (Baylor, 1987). By doing so one can obtain spectral sensitivity curves such as in Figure 2.1.

Decoding the chromaticity and intensity of light The fact that cone types have different peaks of absorption is useful to work out the chromatic property of a light source or an object's reflectance but is not sufficient in itself. Young (1802) and von Helmholtz (1852) recognized that an extra processing stage is necessary to compare the response of the different receptor types. As the photo-pigments absorption spectra are still relatively close to each other, especially between L and M cones, the cone output will be highly correlated. This is why L, M, and S cones are not colour detectors on their own. However, by simple operations between the cone outputs, one can work out chromatic information and intensity. I will describe the processing which happens in the retina, to work out this information as well as its physiological substrate in Section 2.2.3.

Achromatic information (or luminance) can be obtained by summing indiscriminately the three types of cones, this is a measure of light intensity as perceived by the whole

[3] A transducer changes the nature of energy form one type to another.

system. *Chromatic information* can be obtained by subtracting inputs, essentially decorrelating the signals by focusing on the differences between receptor outputs.[4]

Information transfer Photoreceptors do not spike like typical neurons in the cortex but instead undergo modulations of their membrane potential. This potential in turn modulates the transfer of neurotransmitter (glutamate). The synapses between receptors and horizontal or bipolar cells is the second step of visual computation, and are described as post-receptoral. The effect of glutamate on horizontal and bipolar cells has been described by Shiells, Falk, and Naghshineh (1981).

Post-receptoral computation is described in the following section on interneurons and ganglion cells in the retinae, and is the basis of visual pathways. In the rest of this section on cones, I will detail some known cone deficiency to get a grasp of the role that cones are playing in visual processing. Photoreceptors are the first building blocks of perception, I will detail in the next section how some photoreceptors defects can affect perception. This will help understand their primary role.

Photoreceptor deficiency[5] A subset of the population only possesses two cones; these people are called dichromats (e.g., see Kremers et al., 2000). Dichromats have less discrimination for certain colours depending on the cone missing; however, it has been shown that they can detect some colour camouflaged objects better than trichromats (Morgan, Adam, & Mollon, 1992) and display other perceptual advantages (Sharpe, de Luca, Hansen, Jägle, & Gegenfurtner, 2006). Depending on their missing cone type (L or M), more specifically missing opsin photopigment, dichromats are identified as *protanopic* or *deuteranopic*, respectively. The loci of genes, encoding for the pigment responsible for absorbing the light in cones, are on the X-chromosome. Consequently, colour perception deficiencies tend to affect more the male population. However, visual system development appears to be robust and when L or M cones are missing they seem to be replaced on the retina by the other one of the pair (Kremers, Usui, Scholl, & Sharpe, 1999).

The complete absence of a gene type (single gene dichromats) is relatively rare. Mutations in either gene are a more common occurrence and essentially change the peak of spectral sensitivity of the photo-pigment (Jagla, Jägle, Hayashi, Sharpe, & Deeb, 2002; Sharpe et al., 1998). Individuals with this genotype are identified as protanomalous or as deuteranomalous and do not show the same extent of colour discrimination difficulties as protanopes and deuteranopes.

A more severe case called achromatopsia, is a case in which a person cannot distinguish colours. This can be the case if a person does not possess any cones or only the S cone type.[6] Achromatopsia is also associated with other visual deficiencies, i.e. severe loss of

[4]This type of computation is more efficient than having non-overlapping narrow band receptors (essentially the colour detector describe before) as noted by Young (1802), otherwise the visual system would require a detector for each point on the retina for each shade of colour that needs to be detected.

[5]Here meaning, divergence from the more frequent, "normal", genotype.

[6]These individuals are called blue cone monochromats. Note, however, that they retained some residual

acuity.

It is well known lesion studies in neuroscience can tell us about the functionality of specific brain regions. Similarly, it is possible to study the functional role of the cones, by looking into "abnormal" cone phenotypes. From the anomalous visual systems described previously, we can infer two main conclusions.

Firstly, several types of cones (i.e. with different light absorbance peaks) are necessary to perform colour discrimination. By reducing the number of cone types or shifting the absorbance peaks, discrimination is impaired. The reverse statement is also true; extra cones will provide more discrimination power with appropriate neuronal processing. Tetrachromaticty (four cone types) has been hypothesized by de Vries (1948), and is still a research topic of interest (e.g. Jordan, Deeb, Bosten, & Mollon, 2010).

Secondly, the evidence shows that both L and M cone types are necessary for good performance in spatial vision, as acuity is greatly diminished when *both* are missing. However, there is no lack of visual acuity in dichromats compared to trichromats (Jägle, de Luca, Serey, Bach, & Sharpe, 2006).[7] Therefore, L and M cones are crucial for spatial and chromatic perception. A complete loss of S cones (tritanopia) on its own has no effect on visual acuity (D. P. Smith, Cole, & Isaac, 1973), showing that most of the acuity is carried by L and M cone processing. I will now describe the post-receptoral processing in the following sections on intermediate neurons and ganglions cells.

2.2.2 Intermediate neurons

Dowling (1968) identified several types of cells in the retina, here I focus on one subgroup, the intermediate neurons. The intermediate neurons are composed of horizontal cells, amacrine cells, and bipolar cells.

Horizontal cells Horizontal cells and bipolar cells are in direct contact with photoreceptors; through their dendrites effectively forming synapses with the receptors. These synapses communicate signals using of neurotransmitters (glutamate) between cells.

Horizontal cells have dendrites aligned along the retina (hence their name) that sample larger areas of the retina than individual photoreceptors. Horizontal cells are actually modulating receptors at the centre of their *receptive field* (RF). There are possibly three types of horizontal cells (Kolb et al., 1994), named H1, H2, and H3; with different morphological properties and connectivity.

Horizontal cells are averaging the signal over space. They lose the fine grain, as the light information is pooled over a larger area of the cone mosaic, however, detection in the central cones is increased. This is done by lateral inhibition and additionally it has recently been shown that they also use central excitatory feedback in the centre cone (Jackman,

colour-discrimination (Michaelides et al., 2005).

[7] Interestingly, multi-genes dichromats (dichromats with multiple photo-pigment genes, what is referred to earlier as deuteranomalous or protanomalous) showed an increase in acuity compared to trichromats or single-gene carrier dichromats (Jägle et al., 2006).

Table 2.1: Bipolar cell types, IPL: Inner plexiform layer

	Cell Type	
Characteristics	ON	OFF
Light Centre	Depolarize	Polarize
Light Surround	Polarize	Depolarize
Projection	IPL sublayer b	IPL sublayer a

Babai, Chambers, Thoreson, & Kramer, 2011).

Bipolar cells Bipolar cells can be distinguished according to their connectivity (and dendritic tree). They are called bipolar because of their orientation, with the dendrites facing the cones and the axon facing the outer layer of the retina.

There are two main kinds of bipolar cells: midget and parasol. Midget bipolar cells uniquely link to one photoreceptor, as discovered by Polyak (1941). Diffuse bipolar cells connect to a larger number of photoreceptors, around 6 with variations (Boycott & Wässle, 1991). Note that there exist subgroups of these categories with morphological and connectivity differences (e.g. Chan, Martin, Clunas, & Grünert, 2001).

These two types of bipolar cells are divided into two functional sub-types relating to their response pattern. The ON and OFF pathways emerge from these interneurons (Dacey, Packer, Diller, Brainard, & Lee, 2000). Their receptive fields are overlapping centre-surround fields of different polarity and the two populations project to two segregated sub-layers of the retina (see Table 2.1 and Famiglietti & Kolb, 1976; Lennie, 2003; Nelson & Kolb, 2004)

So far, I have shown that retinal processing is dependent on a combination of both horizontal cells (for contrast sharpening) and bipolar cells, through their receptor sampling. Connectivity in the retina after the photoreceptors has been described in detail by Kolb (1970), and more recently reviewed by Lee, Martin, and Grünert (2010). The third type of intermediate interneurons, amacrine cells, seems to perform a similar role as horizontal cells, but at the layer where bipolar cells can be found, their functional role is still debated. I will therefore not discuss this subject further.[8]

The three types of cells I described are referred to as intermediate cells, as they are the step between photoreceptors and ganglion cells, the last type of cells on the retina identified by Dowling (1968).

2.2.3 Ganglion cells

Ganglion cells are responsible for the last stage of retinal processing. They are also responsible for carrying the information to the cortex through the optic nerve. The optic

[8]There is a large numbers of amacrine cell types with associated roles. The interested reader might consult Kolb (1997) for further details.

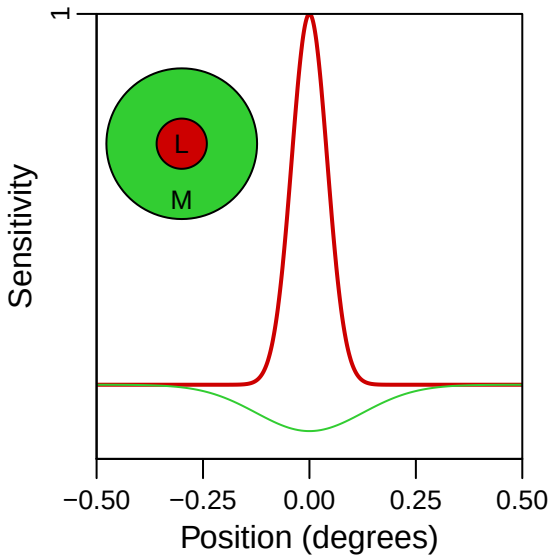


Figure 2.2: Example of difference of Gaussian (DOG) model of a ganglion cell receptive field.

nerve is uniquely composed of ganglion axons. The intermediate neurons, through their horizontal connections and centre surround mechanism, play a role in determining the receptive field of the ganglion cells. The sampling strategy used by the ganglion cells in terms of connectivity (Field et al., 2010) adds to the complexity. The combination of sampling from excitatory and inhibitory interneurons and the centre-surround structure of bipolar cells provide the basis for the shape of ganglion cell receptive fields (Dacey, 2000; Dacey et al., 2000).

Ganglion cells are the precursor of sub-cortical pathways processing colour (S. G. Solomon & Lennie, 2007) and luminance signals. It is notable that the visual channels seem to match the constraints of the retinal level (Lee, 2011). The visual channels are the two colour channels L/M opponent and L,M/S opponent, roughly processing red-green and yellow-blue dimensions and the luminance channel (L+M). The terminology "red-green" and "yellow-blue" is slightly inadequate but widely used. It is required to have both information from L/M and L,M/S cone-opponent channels to obtain colour information (Stockman & Brainard, 2010). These psychophysical channels will be described in the next chapter. They correspond to perceptual sub-organization and are functionally (operationally) defined through psychophysical experiments linked to the type of signal the channels process.

The link between visual pathways, described in this chapter, and visual channels, described in the following chapter, is not straightforward (e.g. Lee, 2011; Tailby, Solomon, Dhruv, & Lennie, 2008) as some pathways carry several post-receptoral signals (double-duty hypothesis, see Section 2.2.3 and Section 3.4.4 p. 52 in the next chapter).

Ganglion cell types

Polyak's (1941) original description of ganglion cells divided them into two groups: midget ganglions and parasol (diffuse) cells. Midget ganglions compute the L-M signal (Dacey & Packer, 2003; Wiesel & Hubel, 1966). Parasol ganglion cells compute the luminance signal by adding several types of cones (Lee, Martin, & Valberg, 1988). Another type of ganglion cell, called bistratified, was discovered by Dacey (1993), and Dacey and Lee (1994). These three types of ganglion cells form the basis of three physiological visual pathways.

Ganglion cells tend to come in two functional types, ON and OFF, responding to the presence or absence of a stimulus. The structure of the receptive field of the retinal ganglion is thought to have a centre (either ON or OFF) and an opposite sign surround (OFF or ON, respectively), see Figure 2.2. This is referred to as connectivity opponency in the rest of this thesis. This receptive field is therefore usually modelled using a *difference of Gaussian* (DOG) model, (e.g. Rodieck & Stone, 1965; Sterling, 1999). However, a recent study by Field et al. (2010), which compared cone inputs to ganglion firing rates in a full retinal map, found that the smooth Gaussian was an inadequate simplification. By using a higher resolution, Field et al. (2010) found that the actual receptive field was point-like at the level of cones, or as the authors described it, "punctate islands of light sensitivity" (p. 673), as opposed to a continuous function (Figure 2.2).

The different types of retinal ganglions and bipolar cells are the basis for the computation, via parallel pathways, of selective visual aspects of the visual scene (for a full review on pathways see Dacey, 2000). The following sections describe the effect of connectivity on ganglion behaviour. The differences in these connectivity patterns are the basis for the differences in visual pathways. Note, however, that a larger number of ganglion cell types, than presented here, has been discovered (Dacey, Peterson, Robinson, & Gamlin, 2003), which might have important role to compute additional features.

Connectivity: Inputs

Midget ganglions have a single centre input from a midget bipolar (with one cone connectivity) (Dacey & Packer, 2003; Kolb & Dekorver, 1991). Parasol cells receive their inputs from diffuse bipolar cells with spatial pooling/averaging (Diller et al., 2004). Bistratified opponents receive an S cone input centre (Dacey & Lee, 1994), and may receive L+M surround input from horizontal cells (Dacey, 2000; Dacey & Packer, 2003).

It is possible to work out the cone inputs of ganglion cells by using stimuli that isolate L and M cones preferentially. Such a methodology was used by Diller et al. (2004) to determine the relative L and M contribution to the midget and parasol cells. These sampled cells were in the peripheral region; therefore, this result might not generalize to the central region of the retina. The authors found that parasol cells summed up L and M inputs, but this was done with an imbalance in L/M input ratio (usually towards L cones). By correlating the results with samples from nearby horizontal cells, they found

that the ratio was determined through the intermediate cells. In turn, the ratio of inputs to the horizontal cells is supposed to be directly linked to the sampling of cones at the retina. This matches results from Dacey et al. (2000), who found that the surround of midget bipolar was probably due to horizontal cells (H1) due to the similar receptive field (RF) size.

In the periphery Diller et al. (2004) also found that midget cells summed up the inputs from L and M cones with similar ratios. They only found one cone opponent midget cell in the periphery. In contrast Field et al. (2010) found numerous colour opponent midget cells in the periphery.

Connectivity: Sampling specificity

Results by Diller et al. (2004) suggest that the variability of the L/M ratio (see Isoluminant Chapter 8) is directly linked to the variability of L/M cone distribution across the retina. The authors showed variation within but also between subjects (macaque monkeys).

In their study, Field et al. (2010) determined the connections for a wide set of ganglion cells. They found that OFF midget cells were five times more likely to sample S cones than the other types of ganglion cells, which can be linked to previous work on the S-cone OFF midget pathway by Klug, Herr, Ngo, Sterling, and Schien (2003). Conversely, parasol cells were never found to sample from S cones in accordance with findings by H. Sun, Smithson, Zaidi, and Lee (2006). The centre/surround opponency arises from a sampling bias (absolute or strong bias) in the centre but random sampling in the surround. Bistratified ganglion cells have been described to have a blue-ON receptive field (Dacey, 1993), but no opponent surround.

Connectivity: Opponency

Opponency, as introduced earlier (e.g. see Figure 2.2), is a property of neurons receiving two opposing signals to compute their outputs, with one being excitatory (or ON) the other one being suppressive (or OFF). Opponency can be, for example, achromatic, i.e. dark versus light, or chromatic, i.e. red versus green (or a composite of both). The opponent process by which ganglion cells extract colour information is referred to as spectral opponency (e.g. Dacey & Packer, 2003).

Opponency is important for spatial computation for two reasons; an achromatic opponency generally computes edges regardless of chromatic properties and is useful for spatial vision. Furthermore, chromatic opponency is crucial to extract chromatic information (S. G. Solomon & Lennie, 2007). As I have discussed in Section 2.2.1, the cone receptors are not sufficient (but are necessary) for colour processing. To extract chromatic information, extra computation is needed (Dacey, 2000). Cone opponency can be obtained by differential sampling of cone types in the centre versus the surround, as hypothesised by Wiesel and Hubel (1966). The midget bipolar cells coupled with midget ganglion cells are a good example for the implementation of this computation because of the single cone

input (Dacey, 2000; Dacey et al., 2000; Lee et al., 2010). Indiscriminate surround modulation can be obtained through horizontal cells (as pointed out by Dacey, 2000) and could be a good explanation for the underlying circuitry (partially confirmed by Diller et al., 2004). Furthermore, as diffuse bipolar cells receive several cone inputs in the centre, it is unlikely that the centre-surround process will generate chromatic sensitivity (Lennie, 2003).

For the case of bistratified ganglion cells, which do not possess surround suppression as pointed out in the previous section, the opponency is within the centre. The receptive field is obtained via S cone bipolar cell (ON) and the inhibition (OFF) is driven through horizontal cells (mainly pooling from L and M cones).

Separation of colour and luminance information

In terms of visual encoding, midget ganglions encode red-green colour opponency, parasol ganglions encode luminance opponency and small bistratified ganglions encode blue-yellow opponency. This section reviews how segregated these different sources of information are.

The retina is the first location where there is evidence of separation of colour and luminance processing (Kaplan & Shapley, 1986). Parry et al. (2012) used a temporal compound stimulus to check the effect of luminance and colour modulation on the ERG (Electroretinogram, electrical activity measured from electrodes placed on the cornea, recording activity indiscriminately from receptors and cells, described in Section 2.2). Their temporal stimulus was designed to separate retinal pathways processing luminance and colour (red-green). They were able to distinguish the two signals showing a differential encoding of both sources of information. However, they did find differences between dichromats and trichromats, which they suggested was due to interactions between the pathways.

Lennie, Pokorny, and Smith (1993) did note the possibility that the luminance function (see next chapter on Psychophysical data) could be computed from linear sums of chromatic responsive cells. Lee, Sun, and Valberg (2011) measured ganglion cell responses to luminance, chromatic (red-green and blue-yellow), and compound (chromatic at spatial frequency f and luminance at 2f) gratings. Ganglion cells (diffuse type) responded more to compound stimuli than single dimension modulations. Luminance encoding ganglion cells did respond similarly to both stimuli. However, the authors did not find any multiplexing (defined below) within the PC pathway. Lennie (2003) suggested that chromatic ganglion cells (and their projecting pathways in the cortex, i.e. the Parvocellular pathway[9]) might encode both colour and luminance at different temporal frequencies.

Multiplexing is a process by which two streams of information are combined into one stream containing both sources of information. It was hypothesized that the red-green channel, starting from the retina, was performing such a task (Billock, 1995; Billock & Tsou, 2004; Ingling & Martinez-Uriegas, 1983). However, the data from Lee et al.

[9]The Parvocellular pathway will be defined in the next section on the LGN.

(2011) suggests strict segregation of the signals (luminance and colour) from the retinal ganglion cells.

The primate's retina is estimated to have 80 different retinal cells varying in several key characteristics. This brief review has outlined our limited knowledge of the full retinal computation. However, some main distinctive groups have been covered. The next section covers the main perceptual relay after the retina, the lateral geniculate nucleus (or LGN) of the thalamus.

2.3 LGN

The LGN is a laminar structure in the thalamus composed of six main layers (plus inter-layers that I will discuss in Section 2.3.2). It is the main relay from the retina to the cortex (Sherman & Guillery, 1996). Its structure is a direct retinotopic map (Polyak, 1957). Furthermore, as described below, it is segregated between layers of cells with different visual properties. The two main layer types are Parvocellular and Magnocellular layers carrying segregated projections from distinct ganglion cell populations.

2.3.1 Parvocellular and Magnocellular neurons

Historically two main types of layers were found, called the Parvocellular and the Magnocellular layers. Wiesel and Hubel (1966) classified neurons from the LGN into four categories. Type I neurons were defined by Wiesel and Hubel as chromatic opponent processing with a centre versus surround receptive field structure. Type II neurons show chromatic opponent processing but no centre-surround structure, i.e. the inhibitory and excitatory regions overlap and these cells do not provide spatial processing. However, type II were not found by Derrington and Lennie (1984).

Type III neurons have a centre-surround organisation with either excitatory or inhibitory centre (and vice versa for the surround) and show no chromatic antagonism. Consequently, type III neurons are thought to carry the luminance information. Interestingly, type III's spectral sensitivity seems to match, broadly, the photopic sensitivity, with a peak at around 550 nm. Type IV neurons showed a more complex pattern, that I will describe later on in this section.

In their study, Wiesel and Hubel (1966) acknowledged that sampling favours the upper layers (Parvocellular[10]) of the LGN. They managed to sample more than 200 cells from these layers compared to about thirty from the deeper layers (Magnocellular). In the Parvocellular layers they found a great majority of cells of type I (164 of 213 cells), mostly of red/green chromatic preferences (but also 3 blue-ON cells, which will be discussed in Section 2.3.2) but also some type III (34/213) luminance responsive cells. Some type II

[10]Note that I refer to Parvocellular and Magnocellular layers throughout this document, however at the time of publication Wiesel and Hubel (1966) used a different anatomical nomenclature which I do not use here for clarity. As a rule throughout this review I try to use modern nomenclature and do not use the terms used in the original papers if those differ.

cells were also found (15/213). From this, we can see that there is a clear preference in the Parvocellular layers for chromatic responsive cells. Derrington and Lennie (1984) confirmed that the Parvocellular layer mostly consists of "spatially and chromatically opponent receptive field (type I)" (p. 237), subdivided into red-green (L-M) and yellow-blue (S- (L+M)).

In the Magnocellular layers the majority of cells found by Wiesel and Hubel (1966) were of type III (21 of 31 cells), the remaining cells they found were defined as type IV and seemed to have complex centre-surround mechanisms, very large *receptive field* (RF) area, and fast non-sustained responsiveness for small spots on the RF. In general, Magnocellular layers were found to respond more to luminance patterns. Derrington and Lennie (1984) abolished the distinction between type III and IV, because they behave similarly to luminance stimuli.

Receptive fields Comparing RF size for cells of type I and III, Wiesel and Hubel (1966) found that type III tended to have a larger RF. Therefore, it appears that there are segregated cell populations with preferential stimulus properties, i.e. chromatic or luminance, and different spatial tuning. Derrington and Lennie (1984) confirmed smaller RFs for Parvocellular (P) than Magnocellular (M) cells. Furthermore, they did not find M cells having RFs in the fovea.

Contrast sensitivity The sensitivity is low for P cells and contrast response is linear, it is also linear for M cells but saturates at higher values (Derrington & Lennie, 1984). The contrast sensitivity, RF, and spectral sensitivity of these cells are highly linked to their pre-synaptic retinal ganglion cells. If these structures can be linked to the different retinal ganglion cell populations, then there is evidence for a physiological basis for an early separation of luminance and chromatic information from retinal ganglion cells through the LGN. Segregated projections forms separated information streams and pathways, I will discuss that point further down in the Section 2.3.3.

2.3.2 The Koniocellular pathway

A third cell population has been identified, more recently, between the Magnocellular and Parvocellular layers, which has been named Koniocellular[11] (see Hendry & Reid, 2000, for a review).

It is now believed that the Koniocellular pathway consists of cells sensitive to Blue-Yellow chromatic information. Recalling some of the data of Wiesel and Hubel (1966), in their cell classification they found a couple of type I cells that they described as Blue-ON centre. Retroactively these cells can possibly be classified as part of the Koniocellular pathway.

[11] From the Greek *konios* meaning dust, in reference to the small soma size of the cells belonging to these intercalated layers.

Koniocellular (K) cells are composed of Blue-ON and Blue-OFF cells, sharing similar response characteristics (Szmajda, Buzás, FitzGibbon, & Martin, 2006, although see Tailby, Solomon, & Lennie, 2008 who showed clear asymmetry between S+ and S- cells), low-pass frequency responses, and large receptive fields (compared to Parvocellular cells).

As with retinal P (Parvocellular) cells (see Lee et al., 2011), K cells respond also to achromatic modulations (Szmajda et al., 2006). However, this is limited to low spatial frequency stimuli, and their response is not as high as for chromatic inputs. This double encoding (or multiplexing) was discussed in Section 2.2.3.

To summarize, the LGN is composed of three main populations of cells. These cells receive projections from the retina and, consequently, project to the primary visual cortex. Cells that project to the visual cortex are called relay cells (65% of all LGN cells; S. G. Solomon, 2002), as they act as a relay from the retina to the cortex, the rest are interneurons. According to the current state of our knowledge of the LGN, relay cells contribute 78% of the cells to the (4) Parvocellular layers , 10% to the (2) Magnocellular layers and 12% to the (6) Koniocellular interlayers, see S. G. Solomon (2002).

2.3.3 Retinal projections

As stated earlier, the LGN is the first relay from the retina on the path to the visual cortex. It is possible to work out the axonal projections from the projection sites in the LGN back to the retinal ganglion cells in the retina by using retrograde tracers. Leventhal, Rodieck, and Dreher (1981), used retrograde tracer[12] injections in Parvocellular and Magnocellular layers to find out which retinal groups projected to those layers. The authors found that Parvocellular cells had projections from midget cells and Magnocellular layers had afferent projections from parasol (diffuse) ganglions.

More recently, Szmajda, Grünert, and Martin (2008) performed a similar experiment in search of the Koniocellular pathway. The injection sites for retrograde tracers tend to leak between layers and therefore are not completely localized to one layer.[13] However, it is possible to determine the extent of the injection sites using histology. The data from Szmajda et al. (2008) elegantly contrasts injection site areas (between layers) and retinal cell distributions between the major known categories of retinal ganglion cells. They reproduced earlier findings showing that the Parvocellular layers mainly connect to Midget cells and a high number of Parasol cells project to the Magnocellular sites. Injection areas with a proportion of Koniocellular cells consistently increased the number of wide field (Bistratified) retinal afferent projections.

Interestingly, Szmajda et al. (2008) found that Koniocellular injection sites consistently had afferent projections from Midgets ganglions. The authors attributed these results to a different set of reasons (such as, leakage from P cells as the pipette goes through layers, easier identification and sampling of midget cells). However, as we have seen in the Retinal

[12] Horseradish peroxide or RPH.

[13] This is because of the difficulty to find the layers and the thickness of the K layer. As stated earlier, this interlayer is thin and composed of cells with small soma.

Section 2.2.3 there exists evidence for Blue-OFF midget cells (see Field et al., 2010; Klug et al., 2003).

The responsivity of the cells in the LGN has been determined to be due more to the direct connection from the retina than from interconnections within the LGN, first suggested by Kaplan and Shapley (1986). Consequently, the pathways are still segregated at this point of the visual processing. Due to this fact, it is now common practice to refer to the midget ganglion cells as P cells (for Parvocellular) and parasol (diffuse) ganglions as M cells. The LGN is thought to conserve the separated ON & OFF ganglion pathways (Nelson & Kolb, 2004), but this segregation disappears in the primary visual cortex.

2.4 Visual cortex

2.4.1 V1

The main projection site of the LGN is visual area 1 (V1) or striate cortex, area 17 in the macaque monkey, where most of the data presented here is from. In this section, I will detail properties of neurons in V1 related to signal processing.

LGN projections V1 has a laminar structure commonly split into six lamina (with sub-laminae). It receives afferent connections from M and P cell mainly to layer 4C where neurons are still monocular (Hubel & Wiesel, 1977). P layer neurons project mostly to layer $4C\beta$, 4A and less so to layer 6 and 1 (Lund, 1988). M layers project to layers $4C\alpha$ and less so to layer 6 (Lund, 1988). K layers (or intercalated layers in the original paper, Lund, 1988) project mostly to cytochrome oxidase-rich blobs in layer 2/3 or in the layer 4A, underneath blobs (Chatterjee & Callaway, 2003; Lund, 1988). Note that P projections in layer 4A seem to project to blobs of layer 3 as well, making blobs uniquely fit to encode chromatic information. Up to this point, it seems that the chromatic encoding pathways (P & K) and the M pathway project to separate layers (Chatterjee & Callaway, 2003). The implications for colour encoding in the cortex are addressed in the Encoding sub-section, p. 17.

Receptive fields properties Neurons in the input layer 4 are still of concentric RF just like LGN neurons; however cells in $4C\beta$ are orientation selective (Livingstone & Hubel, 1984). Furthermore, most cells in other layers show orientation selectivity, with cells of similar orientation being closer to each other (Livingstone & Hubel, 1984). Probing vertically through the layers reveals the same orientation selectivity, thus termed orientation columns (Hubel & Wiesel, 1977).

Orientation selectivity emerges from two types of cells, simple cells and complex cells. Simple cells' RF can be combined by linear sums of excitatory and inhibitory parts whereas complex cells seem to be non-linear (Hubel & Wiesel, 1962). A representation of several dimensions appears to drive or suppress the activity of the cells. The new spatial features

encoded by the visual cortex are orientation and motion direction (Hubel & Wiesel, 1968). Encoding is detailed in the following sub-section.

Excitatory and inhibitory inputs to neurons change the membrane potential of the cell, which will have an effect on the firing rate of the cell and effectively shape its RF by selective sampling. In terms of contrast encoding, the behaviour of a cell related to the membrane potential is non-linear as the firing rate cannot be negative. This is modelled by using a rectification (Carandini, Heeger, & Movshon, 1997), which corresponds to a thresholding of responses.

Cells in layer 2, 3, 5 & 6 start to respond to binocular stimuli (Hubel & Wiesel, 1977), i.e. the RFs of these cells respond to stimuli presented to both eyes (not just monocular as in the LGN), encoding a similar area of the visual field. Cells in the first layer are monocular, probably due to the direct input from LGN neurons. The uppermost layer projects to higher cortical areas (extra-striate): V2 (Livingstone & Hubel, 1984), V3 and V4 (Lund, 1988).

RFs in layer 4 are smaller than in other layers (Hubel & Wiesel, 1977). Importantly, the collective RFs of V1 represent a topographic map of the visual input. Blobs are regularly spaced in this map, therefore they could provide a topographic map of chromatic information (Livingstone & Hubel, 1984).

Chromatic Encoding Initially, in their first report, Hubel and Wiesel (1968) did find some cells (6/25 simple cells and 12/177 complex cells) responding to chromatic components. They suggested recording from "thousands" of cells instead of hundreds to get a better sampling of cortex neurons. Hubel and Wiesel (1968) also described what are now called double opponent cells. Double opponent cells have the opponency present in both centre and surround but with the opponency reversed in the surround, e.g. red-on green-off centre and red-off green-on surround. This type of double opponent cell could be created by sampling from type II cells (with single opponency). Double opponent cells seems to be well suited for texture, form, and pattern encoding, which, Shapley and Hawken (2011) noted, are crucial for object comprehension.

Lennie, Krauskopf, and Sclar (1990) found that single opponent (non-oriented) cells responded to chromatic and luminance stimuli with different spatial frequency tuning. Interestingly, this matches the responses of ganglion (P) cells, which change SF tuning with chromatic components. Therefore, this property might be a leftover of retinal and LGN RF properties. Importantly, Lennie et al. (1990) found that complex cells had similar (band-pass) SF tuning for both luminance and chromatic stimuli. Furthermore, in layers 2/3 & 4, Leventhal, Thompson, Liu, Zhou, and Ault (1995) found simple and complex cells responding to both orientation and colour. There is a multiplexing of colour orientation and edge-polarity in those layers according to Friedman, Zhou, and von der Heydt (2003).

Livingstone and Hubel (1988) described colour properties of blob cells in detail. However, not all colour responsive neurons are blob cells. The authors did find some strong orientation and colour encoding cells in layer 2/3 (between blobs).

The proportion of cells encoding chromatic information is greater than initially thought (Livingstone & Hubel, 1984; Poggio, Baker, Mansfield, Sillito, & Grigg, 1975; Thorell, de Valois, & Albrecht, 1984). Although current research techniques have not yet reached the sampling numbers suggested by Hubel and Wiesel (1968). Johnson, Hawken, and Shapley (2001) found few purely chromatic cells with poor orientation encoding but a large number of colour-luminance cells with strong orientation tuning (30%). Most cells were luminance encoding only (see also S. G. Solomon & Lennie, 2005). The colour-luminance cells were found especially in the layer 2/3 that projects to higher cortical areas, thus shaping the spatial processing of colour and luminance. Those colour luminance encoding cells are mostly double opponent (Johnson, Hawken, & Shapley, 2004), and have a wide range of cone weights giving a distributed encoding of colour edges (see also Lennie et al., 1990, for explanation of that phenomenon see Shapley & Hawken, 2011).

The more recent papers on colour encoding in V1 suggest a ubiquitous presence of colour neurons in V1. This was confirmed with fMRI studies in both humans' and macaques' V1 (Wade, Augath, Logothetis, & Wandell, 2008). The encoding of colour contrast in V1 has been shown to be dependent on the context (Wachtler, Sejnowski, & Albright, 2003), such as surround effects (see Section 2.4.3, 2.4.4 & 2.4.5). Furthermore, Wachtler et al. (2003) showed that population encoding is more representative of perception than single cells. Shapley and Hawken (2011) noted that the visual system might extract hue, saturation, and brightness from a whole population of neurons.

Modular vs. multidimensional representation There is strong evidence relating shape, form, and colour encoding (mainly Friedman et al., 2003; Lennie et al., 1990; Leventhal et al., 1995), reviewed by Shapley and Hawken (2011). Shapley and Hawken (2011) are of the opinion that the modular view of the visual brain (e.g. Livingstone & Hubel, 1988), might be inaccurate. To this effect, they cite psychophysical evidence of cross-masking between psychophysical channels (see my review in the next chapter, Section 3.4). Furthermore, they postulate a greater role of colour than just chromatic computation, i.e. the computation of reflectance vs. illumination (see Shevell & Kingdom, 2008, and next chapter).

The study of spatial interactions and the role played by chromatic information might help to uncover the computational role (as suggested by Shapley & Hawken, 2011) of some of the structures presented in this chapter. The chapter 3, on psychophysical data will emphasise the spatial processing on the computational role of the chromatic and luminance channels.

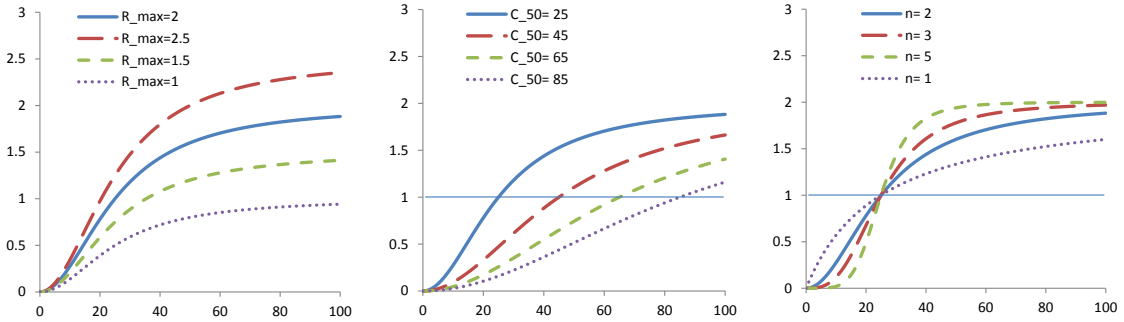
Shapley and Hawken (2011) outline two important ideas, the role of single and opponent cells in the processing of chromatic and spatial-chromatic information and the role of population encoding for hue, saturation, and brightness.

2.4.2 Contrast encoding and gain control

Both simple and complex cells tend to show a compressive behaviour as a function of contrast (Albrecht & Hamilton, 1982). Compressive behaviour means that the response to contrast is not linear but increases then saturates, i.e. it has a sigmoidal shape. It is usually described using the Naka and Rushton (1966) formula:

$$R(C) = R_{max} * (\frac{C^n}{C^n + C_{50}^n}) \tag{2.1}$$

In this equation, C is the contrast input, n is the rate of change, C_{50}^n is the half saturation and R_{max} is the maximum output of the system (see Figure 2.3).



(a) R_{max} determines the maximum output R.

(b) $C_{50}$the semi-saturation constant determines when the function reaches half the maximum values(R_{max}).

(c) n determines the shape of the function making it more or less steep.

Figure 2.3: Description of the Naka-Rushton equation. x axis: contrast. y axis: response strength (here arbitrary unit but can be used to fit spikes per second for example). Note that there is no parameter that forces the function to saturate before a contrast of 100.

The dynamic range is variable between neurons (Albrecht & Hamilton, 1982) so that the high slope will be found at different levels of contrast.[14] Albrecht (1995) found that the semi-saturation contrast tended to be between 10% and 40% contrast (for monkey, V1 cells).

This saturating behaviour is not specific to V1 and can also be found already in the LGN. Sclar, Maunsell, and Lennie (1990) compared firing rates of LGN and V1 neurons and found that the contrast response functions become steeper and saturate faster in the cortex compared to LGN. Sclar et al. (1990) and Albrecht and Hamilton (1982) found that in general the exponent n for simple and complex cells is around 2. Increase in contrast also makes the cell respond faster (Carandini et al., 1997), this can be seen in the EEG too (Burr & Morrone, 1993).

[14] Albrecht (1995) fitted not only the amplitude of the cell responses (in spike/s) but also the phase (in ms) using the Naka-Rushton formula, showing that the phase shift (towards faster responses) also saturated with contrast.

2.4.3 Modulatory effects

DeAngelis, Robson, Ohzawa, and Freeman (1992) and DeAngelis, Freeman, and Ohzawa (1994) reported the distinction between two forms of suppression; one from overlay stimuli and one from purely surround effects. Surround effects (usually suppressive, but not always, e.g. Cavanaugh, Bair, & Movshon, 2002) occur when neuron RFs have a centre with driving excitatory and the surround has no driving effect but shows modulatory effects when the centre is present (DeAngelis et al., 1994, 1992; Hubel & Wiesel, 1968; Jones, Wang, & Sillito, 2002; Seriès, Lorenceau, & Frégnac, 2003). Orientation selective cells will fire when presented with a stimulus in their preferred orientation and do not fire to a cross-oriented stimulus[15] (Carandini et al., 1997; Kimura & Ohzawa, 2009; Morrone, Burr, & Maffei, 1982, but see also Anderson, Carandini, & Ferster, 2000). More importantly, when presented with both stimuli, preferred and cross-oriented, the contrast of the cross-oriented stimulus will have a modulatory effect on the cell response (the common effect reported is suppression; Kimura & Ohzawa, 2009; Morrone et al., 1982). Interestingly D. Xing, Shapley, Hawken, & Ringach, 2005 reported an interaction between orientation selectivity and stimulus size in V1.

The two following sections focus on these modulatory effects, the first one on overlay (cross-orientation) and the second one on modulation due to surround stimuli outside of the classical receptive field.

2.4.4 Cross-orientation effect

Masking suppression can appear when superimposing an extra grating that should otherwise produce no-response, to the preferred grating direction. It is thought to originate from complex cells or a group of simple cells. In some cases, the mask alone is able to activate the cell partially, in which case depending on the target contrast the mask can either enhance or suppress the response. The effect is referred to as crosstalk (Bauman & Bonds, 1991) between spatial channels.

The specificity of cross-orientation is that by itself the cross-oriented stimulus does not elicit any response from the cell. However, once the cell shows increased activity due to the presence of an optimal stimulus an additional cross-oriented stimulus drive the activity down. Priebe and Ferster (2006) studied the mechanism behind this suppressive effect in the cat's visual cortex. The cross-orientation effects are generally thought to be cause by lateral inhibition, after receiving a broadly tuned input from the LGN. Priebe and Ferster (2006) recorded from 32 single cells in Area 17 of the cat[16], using drifting gratings as stimuli. They found that the mask tended to lower the membrane potential. The non-linearity between membrane potential and spike rate amplified this effect and made the cross-orientation suppression bigger.

Interestingly, there was an interaction between test and mask contrast; at low-contrast

[15]The cross-oriented case refers to the orientation relative to the best orientation $\pm 90^{\circ}$.
[16]Equivalent to V1 in primates.

mask (8%) there was no effect on low-test targets (8%) but a slight facilitatory effect on higher target contrast (32% test) with a 5% facilitatory effect (Priebe & Ferster, 2006). The high mask contrast (32%) had far superior suppressing effects. One could expect adaptation effects to play a role in suppressive effects. The authors compared the first and last second of each stimulus presentation and found no significant differences. In physiology studies, cross-orientation suppression has mainly been studied within the *classical receptive field* (CRF). The CRF is the main driver of activity of the cell, although some findings (e.g. Jones et al., 2002), as I will describe in the following section, do not comply with this simple account.

2.4.5 Modulatory effects outside the CRF

Benardete and Kaplan (1999) showed that spatial integration played a role in gain control of M cells, when using large gratings highest contrast would shift the preferred frequency and lower the gain. Comparing small spot and annulus-only responses, it seems that at higher contrast the pattern of activity is more dictated by the surround of the receptive field.

Jones et al. (2002) recorded from 70 cells (simple and complex types), the usual finding in single cell recording is suppression outside the CRF. The CRF is restricted in size and usually when a stimulus goes beyond the CRF, the response starts to be suppressed. This suppression can be attributed to horizontal connection (Seriès et al., 2003) or feedback connections (Angelucci & Bullier, 2003). Usually iso-oriented surrounds tend to suppress the response (from optimal CRF) and for orthogonal surrounds no suppression was found in both single cell recordings (Levitt & Lund, 1997) and fMRI studies (Williams, Singh, & Smith, 2003). Levitt and Lund (1997) did find a predominance of iso-orientation suppression. Interestingly, they did report facilitation of cross-oriented surrounds. These effects were dependent on centre and surround contrast. However, recently Ichida, Schwabe, Bressloff, and Angelucci (2007) found that the far surround could be enhancing. They suggested excitatory and inhibitory regions to be of similar size. This was also found in fMRI studies with dependency on centre-surround relative contrast (Sharifian, Nurminen, & Vanni, 2013; Tajima et al., 2010) and has been linked to a decorrelation of the visual input (Sharifian et al., 2013). Facilitatory effects were also reported in VEP studies (Polat & Norcia, 1998).

Jones et al. (2002) found orientation contrast facilitation in a host of cells. What the authors referred to as orientation contrast facilitation is: facilitation due to a surround differently oriented from the CRF. In their sample of cells, they found five categories. Firstly, non-orientation specific suppression (only 2/70 cells, both S type). Secondly, orientation alignment suppression (12 /70 cells), also called iso-suppression, this suppression diminishes until orientation between centre and surround is ± 45, there was no suppression at higher angles. Thirdly, they found mixed general suppression and orientation alignment suppression. As its name suggests, this category is composed of cells (10/70) showing

both kinds of suppression (described above).[17] Lastly, they described orientation contrast facilitation. This is the most populous group of cells sampled (44/70). These cells showed contrast suppression with iso-oriented surrounds but at higher angles, they showed facilitation, rather than no more suppression. Thus these cells showed more activity (up to +300% or more) when they had a surround at cross-orientation than when the centre was presented on its own. The fifth group of cells is described as orientation contrast suppression (10/70 cells) it is in concept the opposite of the previous group. The suppression from the surround is maximal at cross-like orientation (+-30 °). However, for 8 of these 10 cells, when the separation from inner to outer stimuli was outside the CRF, then the cell switched tuning characteristics to orientation facilitation. This is again an interesting behaviour and partially explains why previous experiments did not find the same results, as inner stimuli size is important.

This paper from Jones et al. (2002) is the first to report such findings. They found the usual cell activity but their new finding showed that V1 cells have a more complex range of spatial processing. Furthermore the response from these surround stimuli was significantly higher than the sum of inner only plus outer only so the effect was not due to a summation of activity. The surround cross-orientation mechanism proposed was a combination of dis-inhibition of the suppressive mechanism and a facilitatory mechanism. This will be interesting to compare with the psychophysics data presented in the next chapter.

2.5 Summary

In this chapter I have presented an introduction to the physiological pathways processing the chromatic and luminance information from the retina to V1.[18] The last sections emphasised spatial responses, however luminance and colour interactions are not emphasized in the literature. The next chapter will present data from the psychophysical approach to perception. The chapter will present psychophysical channels and experiments regarding channel interactions with emphasis on spatial processing and scene understanding.

[17] Effectively, one could model the cells suppression by using a fixed suppression plus a negative Gaussian at iso-orientation.

[18] Further reading: more details regarding light and colorimetry can be found in the Methods chapter, Chapter 4, more details on isoluminance and L/M cone ratios can be found in the Isoluminance chapter, 8.

Chapter 3

Psychophysical evidence for luminance and chromatic channels: independence & interactions

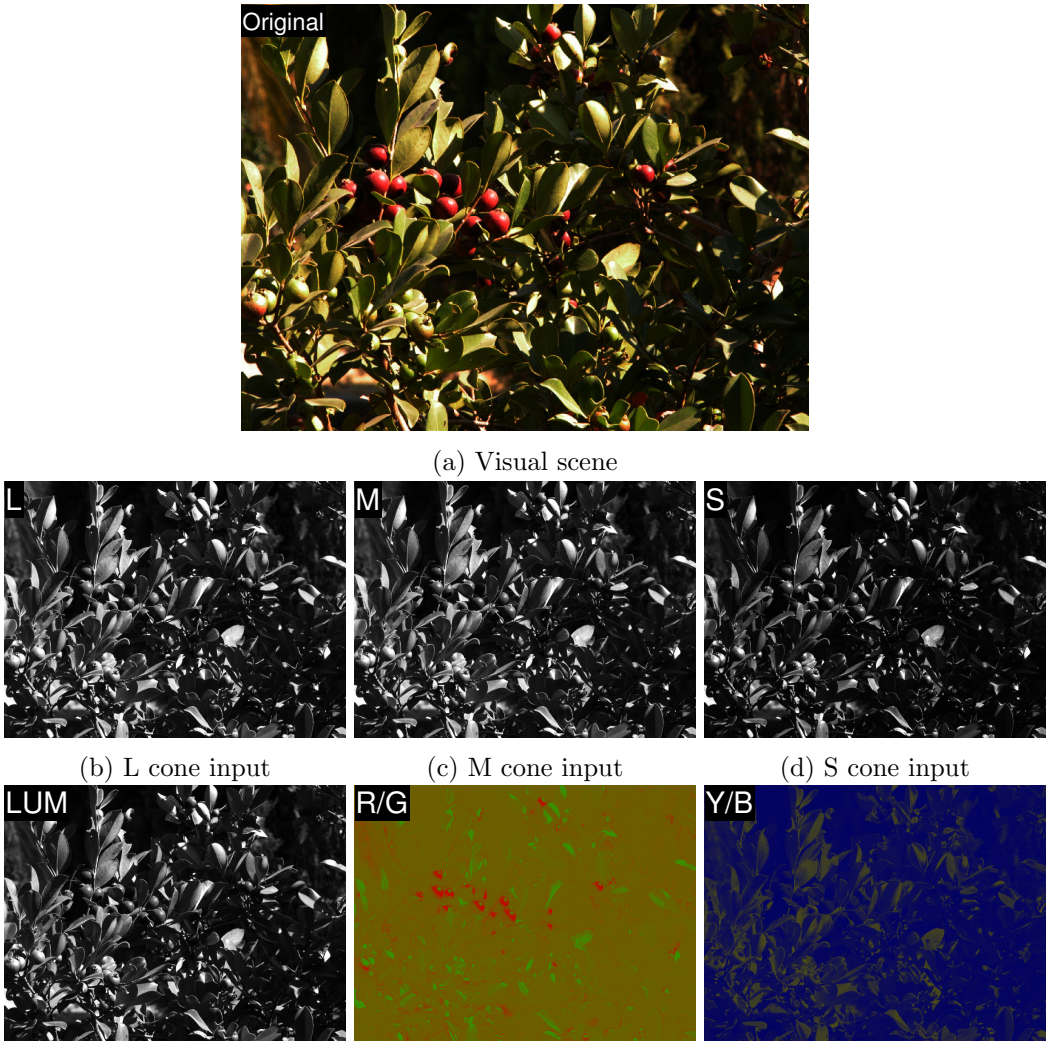
It is thought that the human visual system processes visual information via a set of channels. A channel is a *"functional suborganization within a sensory pathway"* (Regan, 1991). This definition emphasises the role of psychophysical channels, and is based on the results of behavioural experiments. These channels process information in parallel, and the original concept has been attributed (Wilson & Wilkinson, 2004) to auditory work from von Békésy (1960),[1] it was also attributed (see Regan, 1991) to M. D. Wright (1946). M. D. Wright (1946) described the processing of colour as coming from three channels, later to be found in physiology as L, M, and S cones.

From this initial set of channels the visual system forms three opponent channels (see Figure 3.1). The $L + M$[2] also referred to as the luminance channel, the $L - M$ channel or red-green channel and $(L + M) - S$, or yellow-blue channel. This thesis focuses mostly on the luminance and red-green channel. If not specified, when referring to the chromatic channel, the red-green (L-M) is implied.

Thus far, I have briefly described the three main channels conducting spatial processing, however these channels are themselves thought to be composed of narrower channels processing only some characteristics of the visual information. Consequently, I will first introduce the spatial processing characteristics of each channel, obtained through adaptation and masking studies. However, it is now known that these channels do not process information completely independently, consequently spatial channels interact with each other and are pooled across the visual array (Wilson & Wilkinson, 2004). I will therefore describe interaction within channels between spatial frequency and orientation and posi-

[1] von Békésy was awarded a Nobel Prize in 1961 "for his discoveries of the physical mechanism of stimulation within the cochlea".

[2] The contribution from the S cone is limited, but present (Ripamonti, Woo, Crowther, & Stockman, 2009).



(a) Visual scene

(b) L cone input (c) M cone input (d) S cone input

(e) Luminance channel (f) Red/Green channel (g) Yellow/Blue channel

Figure 3.1: Spatial processing stages from original scene to processing by cones (equivalent of retinex pictures Land, 1977) and post-receptoral channels.

tion in the visual field. These types of interactions are the necessary introduction to the core subject of this thesis, the interaction between luminance and chromatic channels. I detail some of the known interactions between luminance and chromatic channels with a variety of paradigms. In the third part of this chapter, I will focus on visual scene interpretation. This section will detail how spatial processing can be used to extract meaningful information from the scene.

In short, in this chapter I will review psychophysical data regarding the spatial processing happening in luminance and red-green chromatic channels. This will take into consideration orientation, spatial frequency and centre-surround modulation effects for detection as well as suprathreshold contrast perception. I will then detail interactions between channels. The first half of the chapter focuses on simple feature detection and is complemented by the second half on information extraction in the visual scene (such as illumination, shape, edges, and objects), which will be detailed in the luminance domain

(as, historically, it was first studied) and with the benefit of colour processing information. The overarching theme is the study of luminance and chromatic spatial channels interactions from edge detection to scene understanding.

3.1 Introduction to luminance and chromatic channels

This section details the general spatial characteristics of the luminance and chromatic channels. The contrast sensitivity characteristics of psychological channels can be inferred by measuring perception within these channels, changing the spatial or temporal characteristics of stimuli systematically. I will focus on spatial and temporal parsing of the visual information. I will also describe adaptation studies in the context of within-channel modulation.

One of the classical methodologies consists of finding the contrast threshold, i.e. minimum contrast necessary for perception. As we are dealing with stochastic processes, a threshold is defined by the experimenter as a certain percentage of detection (usually 50% detection or greater values). The psychometric curve is obtained using, for example, a *two-interval forced choice* (2IFC) or Yes/No paradigm. By changing the spatial or temporal properties of the stimulus, the experimenter can probe what are the best stimuli to drive the channels. In the two sections that follow, I will describe some of the known spatial and temporal characteristics of the luminance and colour channels.

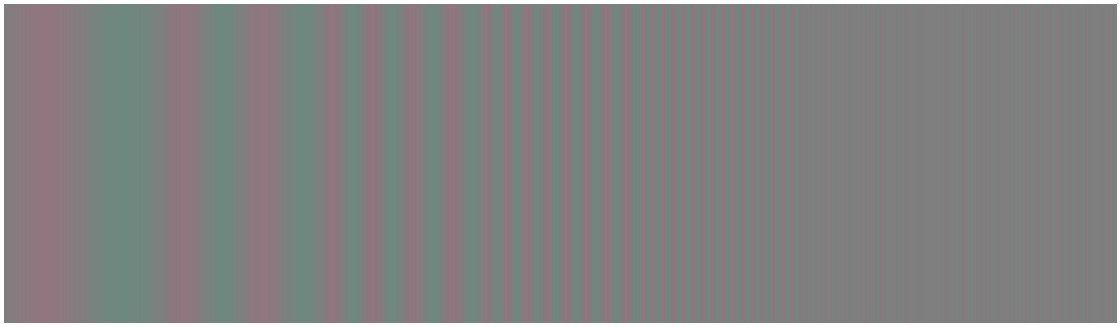
3.1.1 Spatial processing

The spatial characteristics of a channel are obtained by presenting spatially recurrent stimuli, such as a sinusoid[3] or gabor, which vary in *spatial frequency* (SF). SF is commonly described by the *number of cycles per degree of visual angle* (cpd). An illustration of sinusoidal stimuli of varying spatial frequencies and constant contrast is given in Figure 3.2. The spatial frequency of the depicted sinusoid was doubled regularly from left to right. The reader might observe a difference of saliency between luminance and chromatic depending on the SF. This effect is due to differential processing of SF by the two channels.

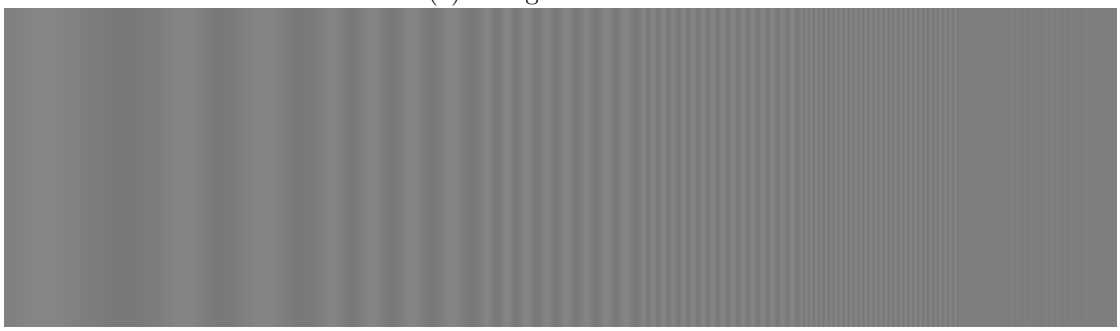
Using a detection task and a range of SF, we can obtain the *contrast sensitivity function* (CSF) of both channels. The usual characteristic CSF of both the luminance channel and the chromatic channel are illustrated in Figure 3.3. The sensitivity is the reciprocal of the detection threshold (1/d) obtained from the detection task. High sensitivity values denote easier detection, given that lower sensitivity requires more contrast to be detected.

In Figure 3.3, we can see that the sensitivity of the red-green chromatic channel (shown by the continuous red line) peaks at low frequency. This low-pass behaviour of the chromatic channel has been found consistently in the literature (DeValois & Switkes, 1983; Gowdy, Stromeyer, & Kronauer, 1999; Lee, 2008; Lee et al., 2011; Losada & Mullen,

[3]The method on how to create sinusoids and equiluminant stimulus is presented in the General Methods Section 4.3.1, p. 85, and in the dedicated Equiluminance Chapter 8, along with contrast, luminance and colour metrics.



(a) Red-green stimulus



(b) Luminance stimulus

Figure 3.2: Illustration of the differential processing of spatial frequency (SF) by luminance and red-green chromatic channels. The stimuli presented have constant contrast. The SF is varied from left to right. (a): the sensitivity is highest at lower SF for colour (low-pass filter). (b): the sensitivity is highest at middle SF (band-pass filter) for luminance.

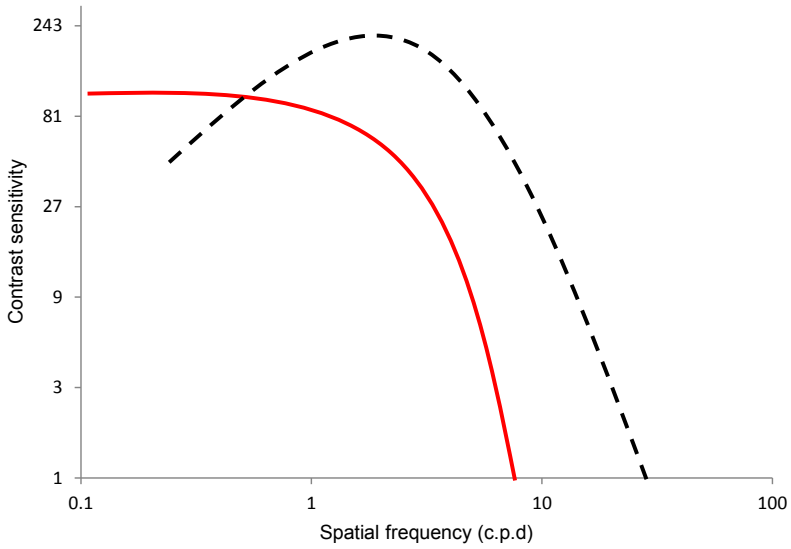


Figure 3.3: Illustration of the contrast sensitivity function (CSF; see text) of luminance and chromatic channels as a function of spatial frequency. Red continuous line: red-green chromatic channel showing a low-pass property. Black dashed line: luminance channel showing a band-pass property at maximum around 2-5 cycles per degree (cpd).

1994; Mullen, 1985; Schade, 1956; Vimal, 1997, 1998a, 1998b; Webster, DeValois, & Switkes, 1990). This has been directly linked (Regan, 1991) to the properties of the cells in the striate cortex (e.g. for cortical cells Thorell et al., 1984; and Lee et al., 2011 for Parvocellular cells showing the two tuning types depending on the type of stimulus[4]).

The luminance channel has band-pass properties, around 2-5 cpd (e.g. DeValois & Switkes, 1983; Wilson, 1990), see illustration in Figure 3.3 (dashed line), and Figure 3.4-a (bold line).

Multiple channel hypothesis

Psychophysical data obtained with adaptation or masking studies points towards the existence of multiple channels (Campbell & Robson, 1968) within the luminance and chromatic sensitivity envelope. An illustration of the multiple channels hypothesis can be in seen in Figure 3.4-a. These channels are obtained by Fourier-transform of the spatial filters presented in Figure 3.4-b. Different functions can be used to model those filters such as *Difference Of Gaussian* (DOG, Rodieck, 1965) or Gabor functions (Marčelja, 1980). These filters were used to describe the receptive field (RF) found in the early visual cortex (see Chapter 2).

The overall shape of the luminance CSF is constructed by a set of finely tuned channels (Figure 3.4-a). The number of channels and their bandwidths is discussed later on.

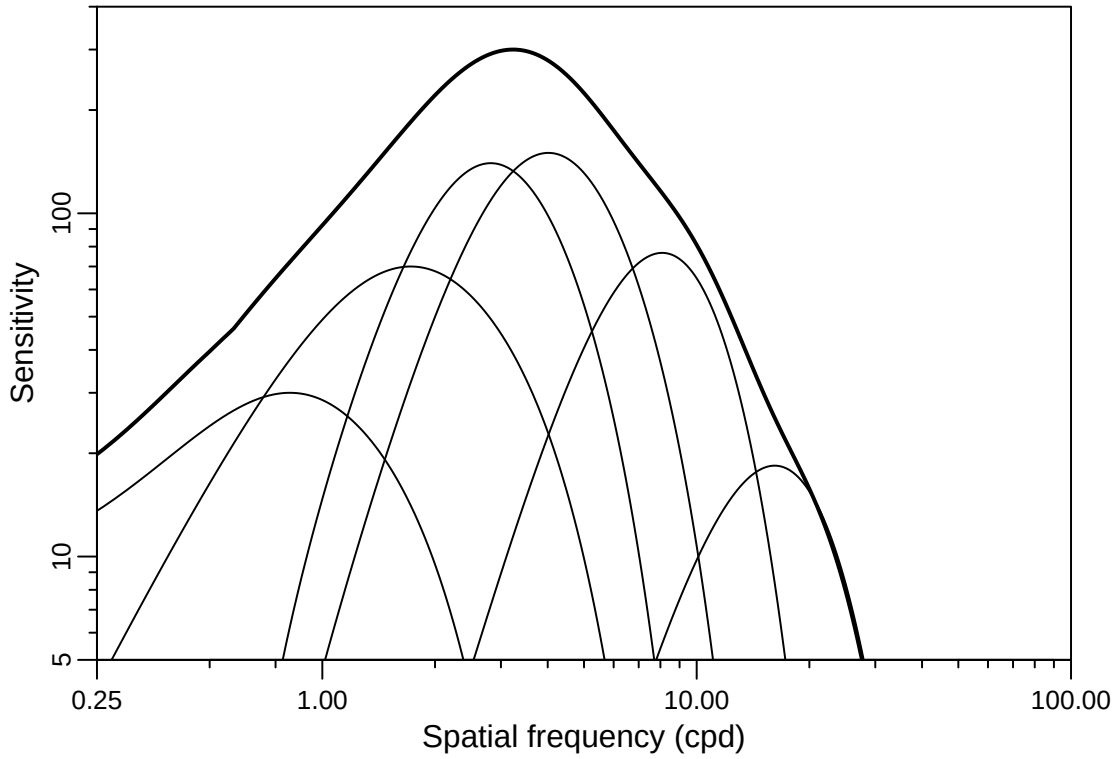
3.1.2 Temporal processing

The equivalent of the CSF (contrast sensitivity function) for temporal tuning is called the Modulation Transfer Function (MTF). It was found, using temporally counterphased sinusoid stimuli (such as in Burr & Morrone, 1993, 1 cpd), that chromatic stimuli are processed by a temporal low-pass filter and luminance by a temporal band-pass filter, see Figure 3.5 for illustration of channel differences. These results have been found consistently in the literature (e.g. Cass, Clifford, Alais, & Spehar, 2009; Metha & Mullen, 1997). These results have also been found using impulse stimuli instead of counterphased, and consequently applying a Fourier-transform, as in Burr and Morrone (1993). It has also been found using both methods by Swanson, Ueno, Smith, and Pokorny (1987). The earliest discovery of the temporal processing discrepancy between luminance and chromatic stimuli is from Regan and Tyler (1971), using uniform flickering fields.

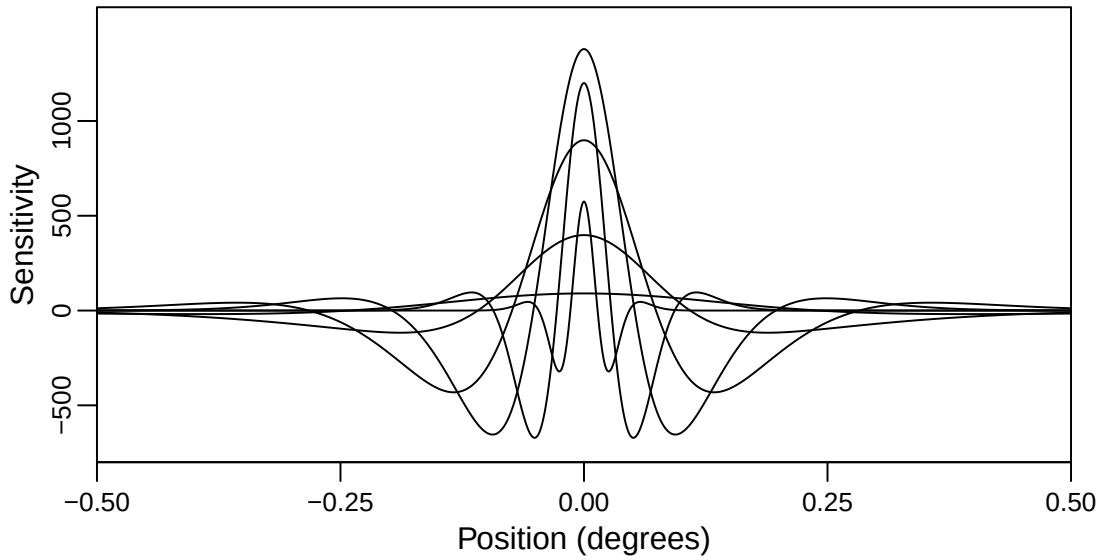
Similarly to spatial processing, temporal processing has been hypothesized to be carried out by several channels. Metha and Mullen (1997) suggested the existence of two temporal channels for luminance and for chromatic processing. For luminance, it was demonstrated that the two temporal channels interact with each other (Cass & Alais, 2006).

Both in this section and in the previous one I have shown one characteristic being studied independently. However, spatial frequency and temporal frequency have been

[4]See also double duty hypothesis in Physiology chapter, Section 2.2.3, see also Billock (1995); Billock and Tsou (2004); Ingling and Martinez-Uriegas (1983).



(a) Multiple channels (thin line) and overall luminance sensitivity (thick line)



(b) Spatial filters for all channels.

Figure 3.4: Illustration of the multiple channel hypothesis. (Filter models and parameters from Wilson, 1990, p69-70)

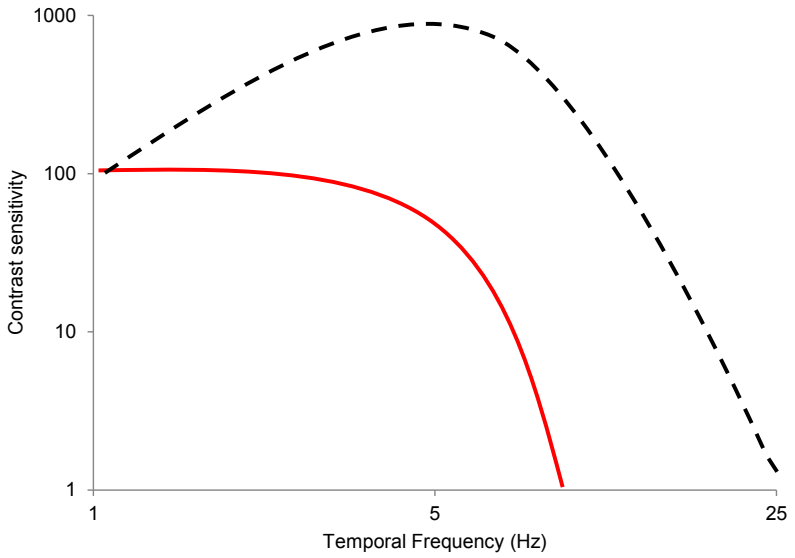


Figure 3.5: Temporal frequency tuning of luminance and chromatic channels. Solid red line: red-green chromatic channel showing low-pass property. Black dashed line: luminance channel showing bandpass property at maximum peak around 8 Hz.

measured systematically by Kelly (1983). This paper showed that spatial and temporal properties of the stimulus interacted with each other and could not be separated by "supposedly" independent functions (for both chromatic and luminance). In a 3D space with the two main axes being SF and cpd and the dependent variable (z-axis) being the observer sensitivity, it was found that the chromatic and luminance sensitivity envelope (or surface) were different. However, Kelly (1983) suggested that the data also provided some evidence for a processing of colour and luminance that might not be completely segregated (see Physiology Chapter 2, Wiesel & Hubel, 1966, Ingling & Martinez-Uriegas, 1983).

In the two next sections, I will detail two main types of experiment to explore in more detail how the spatial processing is performed in the luminance and chromatic channels independently. This will be done through adaptation experiments and masking experiments. Further sections will focus on interactions between channels and their possible use in natural stimuli processing.

3.2 Adaptation studies within channels

3.2.1 Paradigm

Adaptation experiments are based on the observation that long exposure to a stimulus of a specific spatial characteristic (e.g. SF, orientation, position) makes subsequent detection of a similar stimulus more difficult. The effect strength can be modified by changing the contrast, adaptation time and eye(s) used for adaptation/testing (e.g. binocular, monocular or, dichoptic).

3.2.2 Fatigue hypothesis

One of the early hypotheses explaining the adaptation phenomenon claimed that excitation causes fatigue or reduced sensitivity (Ross & Speed, 1991), so the receptive channels (or units) that respond to the stimulus consequently needing more input strength (e.g. contrast or light intensity) to produce a similar output as before adaptation. Adaptation studies can potentially tell us, how stimuli characteristics are encoded together and the extent and specificity of encoding in the perceptual channels as well as the location where this processing takes place.

3.2.3 Multiple adaptation sites

Adaptation can occur at multiple sites: the photoreceptors, the LGN or in a cortical site. The first site is at the receptor level, where molecular adaptation occurs (Hecht, 1937; Stockman, Langendörfer, Smithson, & Sharpe, 2006), which is linked to pigment bleaching. Briefly, at high light intensities the effect of light absorption renders the photopigment (rhodopsin[5] or photopsin) unstable and breaks it into compounds, resulting in bleaching. This process is reversible with time (Hecht, 1937). The bleaching level is not constant and is dependent on prior light levels.

It has been shown that some adaptation effects transfer between eyes (i.e. presentation in one eye and test in the other, Krauskopf, Williams, & Heeley, 1982; Werner, Sharpe, & Zrenner, 2000) suggesting the influence of a cortical site (see Physiology chapter Section 2.4, for binocular processing in the cortex). The third possible site is an intermediary site in the LGN (with only monocular effects).

3.2.4 SF bandwidth

By probing detection continuously with different adaptor SF, adaptation studies can tell us about spatial information processing for the luminance and chromatic channels (Regan, 1991). Studies (e.g. Bradley, Switkes, & Valois, 1988; Ross & Speed, 1991) showed that the highest adaptation effect is at the same frequency as the adaptation stimulus, suggesting a large number of channels (such as in Figure 3.4) or a continuous distribution.

Considering the usual psychometric curve with contrast plotted on the x-axis and detection frequency on the y-axis, the effect of adapting at a different SF than the test seems to shift the resulting slope to the right of the contrast axis; in other words, to obtain similar effect a stronger mask contrast is required (Ross & Speed, 1991). Another way to explain it is that dissimilar SF adaptors have the same effects: a linear increase of detection threshold on log contrast scale, but need equivalently more contrast. The interest is now to find out when this effect disappears i.e. when does the difference in SF become too big that the adapter has no effect on the following detection of the target? We can then conclude that the two signals are processed by separate, non-interacting channels.

[5]The level of intensity for bleaching depends on the photopigment, it already occurs for rods at normal daylight conditions to the extent that they are not usable for daylight perception.

The bandwidth of SF-tuning is measured in *octaves*.[6] The adaptation range was 1.5 octaves for luminance, equally on both sides of the adapter's frequency (Bradley et al., 1988). They found that the adaptation effect was broader for colour by 1.5 octaves and it was also asymmetric. Note that, in that study, the calibration for isoluminance was performed using full-field (i.e. screen) flicker photometric method[7] and, as acknowledged by the authors, might not be adequately tuned to the spatial frequency used (see Anstis & Cavanagh, 1983, for discussion on the subject).

3.2.5 Orientation bandwidth

Similarly to the SF data, Ross and Speed (1991) showed that orientation to a differently oriented stimulus shifts the resulting psychometric function to the right (than iso-oriented); a stronger contrast is required for similar detection levels.

In Bradley et al. (1988), experiments dealing with orientation adaptation (with similar adapter and test spatial frequency but dissimilar adapter vs. test orientation) showed that the adaptation effect worked only in the $[-45°, +45°]$ range, this was the case for luminance and chromatic data. At the time of the publication Bradley et al. (1988) noted that, the physiology and psychophysics mismatched, because no spatially tuned cells had been found for purely isoluminant chromatic information (Livingstone & Hubel, 1984), see also Lennie et al. (1990). However, Bradley et al. (1988) mentioned that some cells were found to encode both luminance *and* colour. These interactions between luminance and chromatic channels are described in a later section, once the characteristics differences between luminance and colour response have been outlined.

To summarize, in this section I have reviewed some of the literature showing some characteristics of the spatial processing of luminance and colour channels with adaptation studies. Most low levels interactions between sinusoidal luminance, or colour, patterns are described in term of masking within channels (e.g. Legge & Foley, 1980, Losada & Mullen, 1994) and between channels (e.g. Chen, Foley, & Brainard, 2000a). In the next section, I present data obtained through masking studies that provides more information about the psychophysical channels. This section will focus on masking within channels (within chromatic or luminance). Interactions between channels (here meaning between luminance and chromatic) will be described in a latter section reviewing masking and other types of studies.

3.3 Masking studies within channels

3.3.1 Paradigm

Masking experiments consist of testing the effect of an additional, irrelevant stimulus (mask) added to a target stimulus that the observer is asked to respond to (e.g. Legge &

[6] In signal processing, an octave is equivalent to a ratio of 2 between two frequencies, f_1 & f_2. It is calculated by $O = \log_2(f_1/f2)$.

[7] Isoluminant methodology is described in Chapter 8.

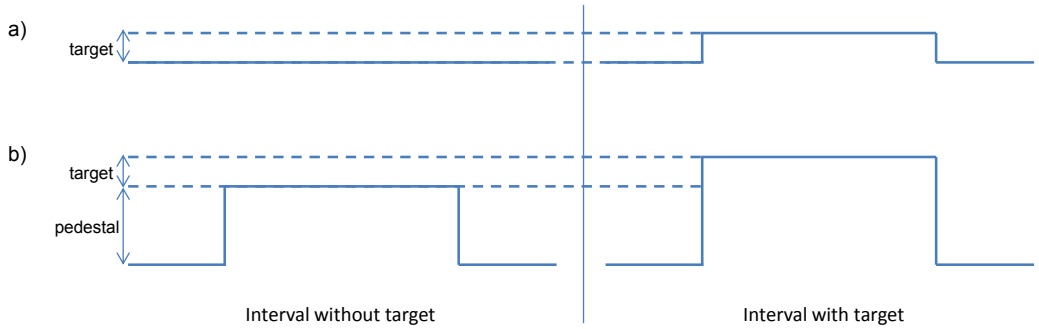


Figure 3.6: Illustrative cross-section of a simplified spatial pattern comparing two intervals in a detection task. The dimension represented shows the target-space. a) Illustration of a classical detection task. In one interval, the target is present and in the other it is absent. The target is defined as a raised signal input from the background values. b) Illustration of a masking paradigm, here the target is also present in one interval but an additional mask is added in both intervals. The aim is to vary the height of the mask to see the effect on target detection.

Foley, 1980). Masking is usually studied to see the effects on the detection threshold of targets, similarly to the adaptation studies mentioned in the previous section. In Section 3.3, I will focus mostly on data collected using a detection task paradigm, and only effects within channel.

The common methodology is to use a 2IFC in which two intervals are presented and in one of them, selected at random, the target is present. Importantly, the mask should be present in both intervals so that the only difference between the two intervals is the presence or not of the target. Figure 3.6 illustrates a detection task with and without mask as a side view of intensity, for a spot-like stimulus.

In the spatial vision literature, sinusoidal modulations of luminance are commonly used for mask and signal, as are gabor functions, however more complex stimuli can be used such as spatiotemporal noise filtering (see Cass & Alais, 2006, Experiment 2). Grating stimuli are also shown in Figure 3.7 showing the two intervals in a detection task, one with the mask only (left of each picture) and one with mask and target (right). Each row shows a mask of increasing contrast from 0 contrast to four times the target contrast.

The mask itself will be processed by the visual system and it is the interaction of processing between the signal (target) plus mask, or mask only, which can help the discrimination between the two intervals. Two important regimes of mask contrast that tell us about the processing (through perceptual channels) are when the mask is below its own detection threshold (or around threshold) and when the mask is really high and discrimination between values becomes difficult (Weber's law). These will be described in detail later on.

To interpret the results of masking studies we need to know how the mask is processed and if the same perceptual channel processes it. The detection of a target (T) by a channel (C) in the presence of a mask (M) only happens when:

$$C(M+T) > C(M) \tag{3.1}$$

This formalised version can be refined to take into account the sensitivity of the mechanism (the amount of extra energy necessary to detect the target in a comparison between two intervals). In the next section, I will describe some of the terminology used in masking studies to describe the type of stimulus, the type of experiment design and the common response curves. I will then present the rationale behind these experiments and the usual result shape from within channel masking studies. The rest of the masking section will then be dedicated to spatial processing (effect of SF, orientation and mask contrast).

Terminology

Before describing the effect of mask on the target detection, we need to differentiate a specific type of mask experiment called pedestal masking from other masking paradigms.

A pedestal is a type of mask "that matches the stimuli in spatial characteristics" according to the definition by Meese, Summers, Holmes, and Wallis (2007). However, the term is sometimes used differently and it is therefore necessary to detail the stimuli precisely when reviewing the literature on the subject. A pedestal experiment using the aforementioned definition can be described as a contrast discrimination task (also mentioned in Meese et al., 2007). For example, one interval will have a gabor of contrast C and another interval will have a gabor with exactly the same spatial characteristics but of contrast $C + \Delta$. In this case it will be evident that the mask (here pedestal) and the target will be processed by the same channel.

However, as mentioned earlier, sometimes the term is used in a more liberal manner, e.g. Chen et al. (2000a) described a "pedestal" that is "tilted at 14.5 re the target" (Vimal, 1997, see also Figure 3.7, right column). The more general definition would be that pedestal and target coincide in space; the spatial properties within (i.e. orientation, spatial frequency) do not have to match. The type of stimulus might be different too. Cole, Stromeyer, and Kronauer (1990) studying luminance and colour interactions (see Section 3.4.3, for full details) using flashes (or spot) of both luminance and chromatic contrast defines a pedestal as being coincident with the target (in time and space) but "irrelevant to the detection task" (p. 128). Masks used in the literature can vary from being identical to the target, to differing in orientation, spatial frequency, temporal frequency, spectral characteristics, or spatial extent (i.e. bigger, or doughnut shaped, see Meese et al., 2007). The spatial relationship between mask and target is important as it can alter the interactions between them (see Figure 3.7, review on spatial effect of relative orientation and extent of masks will be presented in Section 3.3.3 & 3.3.5).

In this thesis, I will use the term mask, and describe the mask as similar to the target or dissimilar, by specifying the characteristics of both. If the term pedestal is used it will be used as defined by Meese et al. (2007).

Regarding the terms used for describing detection task masking experiments, the re-

sults are described as the detection threshold (D) of the target (or the sensitivity, 1/D) against the mask (or pedestal) contrast. The resulting sets of detection levels at different mask levels are referred to in the literature as a masking function or a *TvC function*, or Threshold versus (mask) Contrast. If another parameter is changed, other names are sometimes used such as *TvO*, Threshold versus Orientation (fix mask contrast) and *TvSF* for Spatial Frequency[8], some of these experiments' stimuli are illustrated in Figure 3.8.

Rationale

These experiments are designed to find how detection is altered as the mask is varied in contrast in different spatial combinations (e.g. orthogonal mask, Meese & Holmes, 2007) or with identical target/mask (e.g. Chen et al., 2000a), or both identical mask (pedestal) plus orthogonal mask (producing an overall plaid mask pattern) as in Meese et al. (2007, Experiment 2).

It is also possible to explore the effect of relative orientation of the mask (e.g. Yu, Klein, & Levi, 2001, 2002; Yu & Levi, 2000, for surround masks), spatial frequency (e.g. Legge & Foley, 1980; Swift & Smith, 1983), temporal (Meese & Holmes, 2007), chromaticity (Chen et al., 2000a), on the detection functions. These manipulations are important as they shed light on the underlying channels for spatial, temporal and spectral (luminance and chromatic information) parts of the visual input. Some of these examples are illustrated in Figure 3.8. The results of spatial manipulation will be presented in Sections 3.3.3, 3.3.3 & sec:ch3maskext.

3.3.2 General results: The Dipper Function

Introduction

There is an interesting effect in masking studies that is commonly found and that makes the TvC shape conform to a shape that is called the Dipper function (see Figure 3.9). Dipper functions in masking studies have two components. Firstly, when the mask contrast is low (around or above detection threshold) a facilitatory effect is observed. The facilitation (lowered threshold) is then followed by an increase of the threshold at higher mask contrast.

This pattern of masking occurs within channels; it has been shown for the luminance channel (e.g. Bird, Henning, & Wichmann, 2002; Chen et al., 2000a; Legge & Foley, 1980; Ross & Speed, 1991; Switkes, Bradley, & Devalois, 1988) as well as within chromatic channels (e.g. Chen et al., 2000a; Losada & Mullen, 1994; Switkes et al., 1988), the shape being almost identical (Chen et al., 2000a; Switkes et al., 1988, see Figure 3.9).

Facilitation

Foley (1994) suggested that when the mask was similar to the target it was processed

[8]Note that, as we will see later on, composite experiments are also used sometimes, therefore combining several of these types and exploring a large parameter space.

by the same excitatory units (or channels), which would create the facilitatory effect of the dipper function as the signal plus masker sum up to put the contrast value in the detectable range.[9] This is linked to the non-linear processing of contrast, or non-linear transducer. A transducer is a function relating contrast to detection (in %) as defined in Stromeyer and Klein (1974); Wilson (1980), see Figure 3.10. The transducer function is thought to be non-linear; at some point, the transduction curve gets steeper, which means a small contrast change will have a larger effect on the detection. Consequently, the highest detection gain is obtained near threshold where the slope is expected to be steeper. Similar transducers were found for sinusoid contrast detection (Foley & Legge, 1981; Legge & Foley, 1980; Nachmias & Sansbury, 1974; Ross & Speed, 1991; Speed & Ross, 1992; Stromeyer & Klein, 1974, see also Physiology Chapter 2).

Masking

The masking part of the dipper function has also been hypothesized to arise from the transducer function, when the same channel processes target and mask. The latter part of the function is compressive (see Wilson, 1980 and Figure 3.10), therefore requiring more contrast to detect/discriminate. This is coherent with Weber's law, where higher signal strength delivers higher discrimination threshold (by a constant factor). However it is important to stress that this explanation only works when test and mask are processed by the same perceptual unit (Legge & Foley, 1980; Wilson, 1980) i.e. in the discrimination (pedestal) case.

On the contrary, regarding psychophysical evidence (see Section 3.3.3 & 3.3.5 on spatial processing), Foley (1994) suggested that maskers of all types are summed to produce the suppressing part of the dipper function, as is resultant from broadband inhibition. This new model was required as it explained the broadband tuning of masking in SF and orientation (see next Section) via divisive gain control of broadband origin (Watson & Solomon, 1997).

Tuning of facilitation and suppression

In general, the highest facilitation and masking are obtained when both target and mask are similar and the dipper shape hence tends to occur when the target and mask are spatially identical (Legge & Foley, 1980).

The Foley (1994) and Legge and Foley (1980) studies explored the achromatic channel only. Similar results for the red-green chromatic channel were found by Vimal (1997), i.e. the TvC curves were dependent on the test and mask spatial frequency as well as the orientation between mask and target, and the mask contrasts. However when the mask and target are of different characteristics the resulting TvC function can tell us about the bandwidth of excitatory and inhibitory inputs to a channel. On a more general level, this tells us how the basic building blocks of spatial vision parse the visual information.

[9] This facilitation effect by the masker was referred to as "pedestal effect" in Legge and Foley (1980).

In this section, I have presented a brief overview of the dipper function; for a detailed historical account see J. A. Solomon (2009). I will describe both facilitation and masking in more detail in following sections. I will detail different metrics used to define the facilitation/suppression and estimate their bandwidth in orientation and SF. This will be done for both luminance and chromatic channels. These are an essential introduction to luminance and chromatic channel interaction, which will be studied in the same manner as the within channel-interactions (see Section 3.3).

The complex sets of spatial relationship between target and mask, in term of the TvC shape, are detailed in the next two sections.

3.3.3 Effect of orientation between masker and target

The strength of the suppressing effect for a mask of high contrast (on the suppressing side of the dipper function) is also dependent on the characteristics of both mask and target. This tuning is the main subject of this section. The first half of this section focuses on the luminance channel; the chromatic channel results will be described and compared in the second half. I will first start by detailing some useful terminology.

Terminology

The effects due to target plus mask at a different relative orientation are referred to as *cross-orientation interaction* (XOI). Following Meese and Holmes (2007), I refer to *cross-orientation masking* (or XOM) and the complementary phenomenon *cross-orientation facilitation* (or XOF). This is to distinguish from physiological data from single cells showing *cross-orientation enhancement* (XOE), and *cross-orientation suppression* (XOS), as we have seen in the physiology review (see Section 2.4.4, p. 20). An illustration of the type of stimuli used for these experiments can be seen in Figure 3.8-c.

Cross-orientation sometimes is used in the literature to mean ±90° (orthogonal) or from 45° to 90° (Crook & Eysel, 1992). However, others (e.g. Meese & Holmes, 2007) refer to cross-orientation as any orientation of stimuli that are not aligned (i.e. oblique orientation). In the rest of this thesis I will use this latter nomenclature, with iso-orientation meaning spatially aligned, and will specify the cross-orientation angle used.

Luminance channel

Foley (1994) demonstrated that orientation differences between the target and the masker had an effect mainly on the facilitatory part of the dipper function.[10] The greater the orientation difference, the shallower the dip in the function. At ±45° the facilitation effect disappear completely. This is an interesting parallel to the data obtained with adaptation (see Ross & Speed, 1991). This gives an idea of the specificity of the excitatory units in spatial channels.

[10] Foley (1994) only tested stimuli at 2 cpd.

Masking of a sinusoidal luminance pattern is greatest when the mask and test are about the same orientation (Phillips & Wilson, 1984; Ross & Speed, 1991), but cross-orientation masking does also occur. The relationship between target/mask pair of same spatial frequency and the strength of the masking is complex. The slope of the masking part of the dipper function becomes shallower with mask/target orientation, it reaches its lowest slope when orthogonal (Ross & Speed, 1991). However, Foley (1994) found that the inhibitory effects of luminance masks was almost independent of orientation (only the excitatory input was varied).

There is an interaction effect on the strength of the mask between the orientation and the spatial frequency, in certain SF conditions targets are still masked when masks are oriented at $\pm 90^{\circ}$ (Meese & Holmes, 2010; Phillips & Wilson, 1984). This will be described in the following section on spatial frequency. The rest of this section focuses on orientation effects for colour vision.

Chromatic channel

In colour vision, masking is slightly different: masking appears to be isotropic, i.e. not dependent on mask orientation (Medina & Mullen, 2009). Once again, this has been shown to be dependent on spatial frequency as described by Vimal (1997), which will be detailed in the next section.[11] Vimal (1997) found an orientation tuning that was broader for colour than the one reported by Phillips and Wilson (1984), similarly to the adaptation data.

Orthogonally oriented stimuli usually show masking effects (XOM) but masking can be facilitatory, XOF, (Meese & Holmes, 2007) at low temporal frequency. Medina and Mullen (2009) did find that XOM was independent of temporal frequency but did find XOF in specific conditions (evident in Figure 5 of Medina & Mullen, 2009).

By using monocular and dichoptic stimuli, Gheiratmand, Meese, and Mullen (2013) did find evidence of two separate chromatic channels. One channel is isotropic with contrast summation of orthogonal plaid and monocular. The other one is orientation specific and binocular.

These data focused on the red/green chromatic channel, Chen et al. (2000a) did find that blue-yellow stimuli did not facilitate along the other cardinal colour axis however, non-cardinal signals (e.g. blue-green/orange) did produce the classical dipper function shape. Cross-channel interactions will be detailed in a latter section.

Summary

As a rule of thumb, the more similar the mask and target are the greater the masking effect. This is because it is thought that the two are processed by the same channel and the mask is adding noise rendering detection more difficult. Similarly, facilitation is greater at the same spatial frequency (as we will see in the next section).

[11] Vimal (1997) showed at high mask contrast values that this isotropic is only present around 0.5 cpd.

However, the tuning is different between facilitation (sharply tuned) and masking (broadband). These differences might suggest that facilitation and masking are due to two different mechanisms (Legge & Foley, 1980; Mullen & Losada, 1994; Ross & Speed, 1991).

Chromatic XOM is stronger than luminance XOM (Medina & Mullen, 2009; Mullen & Losada, 1994), this difference has been attributed to stronger masking at a monocular stage (Kim, Gheiratmand, & Mullen, 2013). XOM is independent of temporal frequency for the chromatic channel (Medina & Mullen, 2009) but this is not the case for the luminance channel (Medina & Mullen, 2009; Meese & Holmes, 2007).

The next section will focus on the SF tuning of interactions within luminance and chromatic channels.

3.3.4 Effect of spatial frequency

In this section, first I will review studies testing target and mask with different SF to look for the SF bandwidth of facilitation and suppressive effects. I will then compare the absolute facilitation and suppressive effect values when mask and target have similar SF, to search for a main effect of SF. I will then compare the luminance and colour channels, and finally describe possible interactions between orientation and SF.

TvSF function

When target SF is fixed and mask SF is varied (TvSF, see Figure 3.8-b), the highest masking effectiveness occurred when mask and target were at a similar spatial frequency. This was found for the luminance channel (DeValois & Switkes, 1983; Legge & Foley, 1980; Losada & Mullen, 1994; Speed & Ross, 1992; Switkes et al., 1988) but also for the chromatic channel (DeValois & Switkes, 1983; Losada & Mullen, 1994; Switkes et al., 1988; Vimal, 1998b).

DeValois and Switkes's (1983) data was obtained using a fixed mask contrast (20%) and a wide range of SF, similarly to Vimal (1998b, Experiment 1, threshold elevation) at 60% contrast. However, the other papers mentioned used designs with full TvC functions. This consists of testing the full contrast curve (TvC design, e.g. Figure 3.8-a) for a range of mask SF (e.g. Figure 3.8-b).

Estimation of SF bandwidth Here, I am interested in the SF bandwidth of the facilitatory and suppressive effects. These two tunings will be described in the two following subsections.

Bandwidth estimates depend on the method used to obtain the data. This is so because contrast has a suppressive effect. Consequently, bandwidth estimates using the full TvC functions (Legge & Foley, 1980) are more appropriate. Furthermore the use of different metrics, on the full TvC, will have effects on the bandwidth estimate too (Losada &

Mullen, 1994; Speed & Ross, 1992; Switkes et al., 1988).[12]

Several studies that studied facilitation and masking SF bandwidth of masking are presented below.

Facilitation bandwidth The facilitatory dip appeared at lower mask contrasts for masks/targets of similar SF, and moved toward higher contrast values when the mask was away (in SF space) from the target (Legge & Foley, 1980; Speed & Ross, 1992).

The magnitude of the dip was greatly reduced away from the target SF in Legge and Foley (1980). The bandwidth estimate was ± 1 octaves for Legge and Foley (1980), ± 0.5 octaves for Switkes et al. (1988) and ± 1.5 octaves for Speed and Ross (1992), using a different metric.

For the chromatic channel, the facilitation was not found consistently in Vimal (1997) but the design of the TvC experiments did not sample the contrast space as much as Legge and Foley (1980). With a more thorough sampling, Losada and Mullen (1994) did find facilitation consistently in the chromatic TvC function.

Switkes et al. (1988) found exactly the same SF tuning for the chromatic channel, confirmed for the luminance channel and chromatic channels by Losada and Mullen (1994). Losada and Mullen (1994) noted that the facilitation amplitude was prone to individual differences (already noticeable with 3 participants). This is a point I will come back to when describing my own experiments.

The facilitation results are important in the light of the spatial channels presented earlier, as it seems that the facilitatory effect bandwidth is around ± 1.5 octaves, which is somewhat similar to adaptation bandwidth effects. This is not surprising as both phenomena are supposed to be linked to the excitatory drive, i.e. the excitatory input that drives a perceptual channel response.[13]

Suppressive bandwidth Contrast has an effect on the suppression strength and SF bandwidth. The SF tuning of suppressive effects becomes broader with contrast, after the facilitation dip (Legge & Foley, 1980). Furthermore, the metric used to define suppression has an effect (see Section 3.3.4, Estimation of SF bandwidth).

Masking was found to be highest at the point where target and mask are the same (e.g. Losada & Mullen, 1994), as it is the case for the adaptation literature, and for masking facilitation. However the SF tuning of suppression is broader than the tuning for facilitation (e.g. Switkes et al., 1988).

Luminance suppression is asymmetrical with shallower falloff at higher frequency than lower ones (DeValois & Switkes, 1983; Legge & Foley, 1980; Switkes et al., 1988) and

[12] Switkes et al. (1988) used equated suppressive effects (in threshold elevation ratio) to compare luminance and chromatic channels. Losada and Mullen (1994) used the turn of the facilitation dip as criterion for suppression.

[13] Also Bradley et al. (1988) found that adaptation strength effect, the suppressive effect of an adaptive field, is dependent on its contrast. The overall orientation spread and magnitude are increased by contrast. Here when the SF is different more contrast is needed to reach the facilitation dip, making it also shallower. See Ross and Speed (1991) for comparison between the two paradigms.

this asymmetry gets stronger with contrast. At higher mask contrast the higher SFs are more involved in the suppressive patterns, this asymmetry seems to start between 6.4 & 12.8% luminance contrast mask (in Legge & Foley, 1980). This asymmetry might be due to the CSF function (see Section 3.1.1).

For the chromatic channel, DeValois and Switkes (1983) found highly tuned suppressive effects of chromatic target by chromatic mask, which they argue is a sign of narrowly tuned spatial channels within the chromatic channels (see Introduction on multi-channel hypothesis, Section 3.1.1). They found that the chromatic masks were more efficient at masking. However, when luminance and chromatic masks are equated for potency, chromatic masking is only slightly stronger in strength (Switkes et al., 1988). It also appears only slightly more tuned, both channels have a suppression bandwidth around ± 2 octaves. The cross-channel interaction case will be discussed later.

Note that all of the data mentioned above is for iso-oriented conditions.

Main effect of Spatial frequency

The experiments described here involved mask/target pairs of similar SF. This section is focused on the magnitude of facilitation and suppression at different SF.

Facilitation Speed and Ross (1992) found that facilitation was highest at low frequencies then dropped after 5 cpd, by a 100 fold magnitude at 20 cpd. There is still a small facilitation effect visible at 20 cpd but it occurs at high mask contrast values. Losada and Mullen (1994) found that the amount of facilitation and its tuning seems to be dependent on the participants (testing only low SF).

Suppression In the luminance domain, Phillips and Wilson (1984) showed a main effect of SF (in transient presentation). The tendency is that lower frequencies have higher suppressive masking. This effect is also visible in Meese and Holmes (2007) using the same range of spatial frequency, but using orthogonal masks.

However, Losada and Mullen (1994) tested a lower range of SF (0.25, 0.5, 1 cpd) and found a marginally higher masking effect at 1 cpd for luminance, but constant or decreasing chromatic masking. Masking is higher at really low frequencies for colour compared to luminance (Losada & Mullen, 1994).

Switkes et al. (1988) demonstrated a constant suppressing power for colour (from 0.125 to 4 cpd), for luminance, it is almost constant over the range of frequencies. However, slight tuning is visible, peaking at 2 cpd (see Figure 9-a open circles in original paper). Switkes et al. (1988) showed that masking is more potent for the chromatic channel compared to the luminance channel at really low SF. Vimal (1998b) found that at 8 cpd the masking was completely abolished for one participant and very diminished for another one.

Data from Meese and Holmes (2007), for the sustained case (referred to as 0.5Hz condition, in their paper) showed similar masking at 0.5 and 1 cpd and then decreasing masking effects (**0.5**, **1**, 2, 4, 8 cpd). In summary, for luminance the suppressing abilities

tend to be constant for low spatial frequencies then drop around 10 cpd or before. This appears to be faster for the chromatic channel (Vimal, 2002).

Summary of differences between luminance and chromatic channels

They are some common characteristics between chromatic and luminance masking. The Dipper function shape has been found for luminance and chromatic channels, the facilitation point is always around the detection threshold (e.g. Switkes et al., 1988). Using threshold normalized data, Switkes et al. (1988) found that the two dipper functions aligned (this was tested both at 2 cpd and 0.5 cpd) and this was confirmed by Chen et al. (2000a), see Figure 3.9, p70.

Chromatic XOM appears to be stronger and broader. As in adaptation studies (e.g. Bradley et al., 1988), the SF tuning of chromatic suppression is broader than chromatic tuning (also found in Switkes et al., 1988). Further, this masking was found to be symmetrical for chromatic and asymmetrical for luminance (Losada & Mullen, 1994; Switkes et al., 1988).

Interestingly, Vimal (1997, 1998b) showed that the suppression effect was larger for lower spatial frequencies for the red-green chromatic channel. Comparing Phillips and Wilson (1984) and Vimal (1997)'s data, Vimal (1997) found that achromatic and chromatic differences were smallest at 2 cpd (in accordance with Switkes et al., 1988), at the peak of the achromatic CSF.

DeValois and Switkes (1983), found that both colour and luminance suppression were SF tuned but chromatic was more sharply tuned. In terms of absolute amplitude of the suppressive effect, when masks are defined as multiples of detection threshold the suppressive effects are of similar magnitude.

Interaction between orientation and SF?

As mentioned previously, there is an interaction between mask orientation and SF. We have seen that mask's orientation and SF away from the target lower the strength of the masking/facilitation (Ross & Speed, 1991). In same SF pairs, we have also seen that there is a tuning of the suppressing function.

At 8 cpd, Phillips and Wilson (1984) found that the masking becomes weaker and when the orientation reaches 45 degrees, it becomes completely abolished. Very similar results were obtained by Vimal (1997) for chromatic stimulus. Although Medina and Mullen (2009) showed that XOM is isotropic and non-interacting with SF. These two results imply a tightening of the orientation bandwidth of suppression for high SF, equally for luminance and chromatic channels.

Meese and Holmes (2010) confirmed that XOM was stronger at lower spatial frequency, however they also found that the interaction with orientation should be re-evaluated and might be due to different processes. They argue that the orientation bandwidth is constant around $\pm 20°$. Meese and Holmes (2010) hypothesized that there are at least two

suppressive processes. One process is strongly orientation selective, of cortical origin, and the second is weaker, isotropic, but can get stronger suppressive power at low spatial frequency or high temporal frequency. This would also fit the chromatic suppression data by Vimal (1997). Interestingly for the chromatic channel, there is also suggestion of two tuned suppressive mechanisms (Gheiratmand et al., 2013).

Meese and Holmes (2010) noted that other factors could influence suppression such as tuned interaction or surround effects.[14] The surround effects are going to be described in the next section.

3.3.5 Effect of spatial extent of masks

They are a small set of mask and target spatial configurations that have been studied extensively. In spatial vision, the targets used are generally circular patches made with a sinusoidal grating (or gabors). Target extent (without any mask present), will have an effect on detectability due to spatial pooling. In masking studies, the masks used can be of various extents (similar, extending beyond or surround only.[15] Surround only (doughnut shaped) masks are used usually because of the parallels with physiology (see Section 2.2.3 in Physiology Chapter 2).

Foley (1994) used both extended mask (full field) and a local gabor mask and found greater facilitation with the gabor than the full field mask. Yu and Levi (2000) and Yu et al. (2002) showed that surround could facilitate detection especially in cross-condition, which lead to their proposal of two surround mechanisms (Yu & Levi, 2000, although see Meese, Challinor, Summers, & Baker, 2009). Additionally, Petrov, Carandini, and McKee (2005) demonstrated the existence of different suppression mechanisms, the overlay suppression and the surround suppression, echoing the physiological data presented above. Surround suppression is more tuned (for SF & orientation) than the overlay type, furthermore it gets stronger with eccentricity (Petrov et al., 2005), saturating at around 4 deg of eccentricity. This indicates more localized spatial processing in the fovea compared to the periphery. (Petrov et al., 2005) Overlay masking precedes surround effect and surround masking saturates overlay does not (Meese et al., 2009; Petrov et al., 2005).[16]

This difference in processing between centre and surround was explored in more detail by Meese et al. (2007). In their dipper function model (based on the original work by Foley, 1994) they include specific parameters for surround and centre modulation, with different weight of suppression and facilitation (hence excitation and inhibition) in the centre and the surround. These weights varied between participants. Additionally, Meese et al. (2007) performed experiments with the three types of masks mentioned. They concluded that both excitatory and inhibitory effects existed for cross-oriented stimuli in the centre as well as in the periphery. The role, extent and physiological reasons for these multiple interactions are still debated.

[14] These effects can alter the proper estimation of filter bandwidth (see Section 3.1.1, on filters).
[15] Or surround only with gap.
[16] This interacts with orientation (Meese et al., 2009).

3.3.6 Summary of within channel masking

In Section 3.3, I have shown that the processing of spatial information is performed by both the luminance channel, as expected, but also significantly by the chromatic channel (e.g. Werner, 2003). This has some important ecological significance in terms of extraction of scene information such as reflectance, outline, and shape of objects where the chromatic information can play an important role (see section on Scene parsing 3.6).

Furthermore, I provided evidence for the existence of both facilitatory and suppressing processes with specific tuning properties. The suppressive mechanism might be composed of several sub-mechanisms (Gheiratmand et al., 2013; Meese & Holmes, 2010; Petrov et al., 2005). The usual shape of within channel interactions is the dipper shape function (Chen et al., 2000a), probably resulting from the compressive nature of the channel (Legge & Foley, 1980). Both the target and the mask spatial characteristics will influence the shape of the TvC, and by studying these characteristics we can learn how spatial information is encoded within a channel. The excitatory input seems to be tuned and the suppressive drive more broadband, yet the overall number of mechanisms within suppression is still debated (Meese et al., 2009; Meese & Holmes, 2010).

The link between psychophysics and physiology is promising and has long been recognized in the masking literature; however, we must be careful on bridging the results between (V1) single cells and perception. At the systems level one observer is able to draw inputs from a pool of detectors (see for example Meese & Summers, 2012 on area summation and detection). A complete understanding of spatial channels requires understanding of interactions within channels, spatial interactions (such as spatial pooling, centre/surround), but also interactions between channels and this must be known for detection but also suprathreshold perception.

In this section, I have described some of the properties of the luminance and colour channels obtained through masking studies (at detection level) and described some of the spatial integrations (surround effects). In the next section, I will focus on interactions between these channels.

3.4 Luminance and chromatic interactions

Vimal (1998b) pointed out that an understanding of the interactions between luminance and colour in contrast processing was still "largely unclear". This same question was raised by DeValois and Switkes (1983) whose work was one of the earliest attempts to link the two channels.

In this section, I will describe the data obtained with adaptation, masking and flash detection regarding luminance and chromatic (red-green) interactions. I will then detail some of the computational theories put forward to explain the data.

3.4.1 Adaptation data

Cross-adaptation

Bradley et al. (1988) demonstrated cross-channel adaptation, using red-green adapter and luminance targets (iso-oriented). The extent of adaptation effects was not as high as for luminance adaptors on luminance targets. For the reverse condition, luminance adapters and colour targets (iso-oriented), Bradley et al. (1988) found zero (at 1 cpd) to very modest adaptation (at 2 cpd).

Lack of cross-adaptation effects was previously found by Krauskopf et al. (1982), they also tried mixed adaptation (exciting more than one of the three main channels, i.e. mixed luminance, red-green or yellow-blue), and found a less chromatically tuned adaptation than with cardinal adaptation showing evidence of a three axis colour space.[17]

Chromatic adaptation & luminance structure

Adapting to a chromatic background can subsequently change the perception of other subsequently colours presented, through a process called *chromatic adaptation*. Chromatic adaptation as a perceptual phenomenon has been linked to colour constancy. Concretely, the effect is a shift of the white (or grey) point in the colour space[17] towards the adapter colour (see Webster, 1996, for an extensive review on colour and adaptation). To simplify, if an observer adapts to a green light environment the colour green will eventually look grey/white. This effect has been attributed a functional role that discounts the illuminant to work out the real reflectance of objects, which is important for colour constancy (Webster, 1996). However, the effect strength is dependent on the spectral composition of the adaptive stimulus (Werner et al., 2000). In this section, I present results related to the specific interaction of luminance and chromatic adaptation phenomena.

Using a honeycomb pattern (see Figure 3.11-a,b) surrounding the test patch (as opposed to uniform), Werner et al. (2000) demonstrated an increased chromatic adaptation (i.e. increased shift in white point compared to using a uniform background, Figure 3.11-e,f). This enhancement was present if the pattern was luminance, chromatic or luminochromatic defined. Furthermore, for luminance, the contrast of the pattern increased the adaptation effect.[18] This luminance specificity of chromatic adaptation is akin to the McCollough effect, which will be described in the following section.

When the surround was composed of stripes (see Figure 3.11-c,d) of different SF the adaptation shift[19] peaked when the surround SF matched the test size. The adaptation decreases down more sharply for lower background SF than higher; it is interesting to compare this data with TvSF functions displaying similar behaviour. Orientation effects have also been found using a bar test and stripe background of different relative orientation (Werner, 2003). Adaptation was highest when surround and test or adapter and test had

[17] Refer to the Methods Chapter, Section 4.2.5 for more details on colour spaces.
[18] This luminance effect saturates at 10% contrast (Werner, 2003).
[19] The adaptation shift corresponds to the shift of the white point locus on a given colour space.

the same orientation. These spatial interactions are similar to adaptation and masking luminance detection tasks, described in earlier sections.

Adaptation is therefore spatially dependent, and furthermore a purely luminance background structure can modulate this adaptation (Werner et al., 2000) which highlights the interaction between chromatic perception and luminance content. Dichoptic viewing yielded larger effects than binocular and the effect did transfer between eyes without loss of potency (Werner et al., 2000). This provides evidence of a cortical origin for these interactions. Werner et al. (2000) also found differential processing speed of adaptation depending on the colour. Yellow and green were faster and more potent than blue and red backgrounds. This difference only exists when the adapter/test pattern is a complex lumino/chromatic structure. Using a uniform background the processing time (and magnitude) of the adaptation effects are similar.[20] The luminance structure had no effect for red and blue adaptation (in overall magnitude and speed).

Using textured backgrounds, Zaidi, Spehar, and DeBonet (1997) found that spatial structure (chromatic and SF variations) could mask detection of chromatic changes. Consequently, textured background can deliver a complex set of interactions with colour perception.

In this section, I have shown that colour perception and adaptation is modulated by spatial content over a quick time scale. The next section on the McCollough effect will describe a similar chromospatial phenomenon that acts over a longer period.

McCollough effect The McCollough effect (McCollough, 1965), is a long lasting adaptive effect that shifts the chromatic content of achromatic stripes. The effect is contingent on the adapter and test orientations.

The adapter usually consists of two time-interleaved adapters; one set of red/dark stripes and one set of green/dark stripes. One of the adapters is horizontal, the other vertical. When a target of similar orientation is presented, its chromaticity will be shifted away from the adapter colour. For example, if the adapter were horizontal green, an achromatic horizontal set of stripes would appear reddish (and vice versa for vertical red adapter). This demonstrates a dual encoding of orientation and chromaticity.

As is often the case with adaptation effects, the perception of the test is away from the adapter direction, but here the post-adaptation effect is contingent, or bound, to the target orientation. Hence, the adaptation is described as separable (see also tilt after-effect Section 3.5.2).

The specificity of adapter orientation, luminance and chromatic contrast is the main driver for chromatic effects on the adapters (Webster & Malkoc, 2000). Higher target contrast produces higher chromatic shifts; furthermore, adaptation to purely isoluminant stimuli strongly extinguished the effect. Webster and Malkoc (2000) pointed out that this suggests an aftereffect based on lumino-chromatic interactions with orientation encoding. The phase of the luminance/chromatic relationship in the adapter only changes the mag-

[20] See also Krauskopf et al. (1982) on differences between red and green adaptation.

nitude of the effects on the adapted bars but does not extinguish it. Lastly, Robinson and MacLeod (2011) showed that the effect works using plaid adapters and showed dissociation of effects for grating and test plaids. They suggest that a plaid encoding might exist on top of the grating processing. This encoding is an extra-step toward the representation of complex objects or textures.

Summary The data presented here shows that there are complex interactions between chromatic appearance and spatial information encoding. Webster and Malkoc (2000) suggested that colour-luminance contrast is encoded together or distributively encoded over a population of neurons and that the non-interaction of isoluminant adapters reflects the asymmetry in colour and luminance encoding. The interaction asymmetry is an important part of the masking results presented in the following section and will be discussed further there.

3.4.2 Masking data

This section focuses on colour-luminance interactions in the context of masking using a detection task paradigm. In practice, I will be reviewing results from experiments using colour & luminance mask/target pairs in the detection paradigm. These studies focus on the independent detection of either colour or luminance variations in the presence of the other type of stimulus (as a mask). The two combinations will be explored, and as in previous sections on masking, SF and orientation tuning will be detailed.

Colour Mask/Luminance Target

In this section, I will consider cases where luminance targets are detected when a chromatic (red-green) mask is present.

Fixed contrast results Using fixed mask contrast, colour masks were found to suppress luminance targets (for iso-oriented 20% contrast DeValois & Switkes, 1983, and +14.5° orientation and contrast of 2.7 times the threshold in Vimal, 1998a). DeValois and Switkes (1983) found that the suppressive effect was similar to the one obtained from luminance masks (of 20%), confirmed by Vimal (1998a) using the appropriate multiple of threshold contrast metrics. Compared to within-channel masking, higher masking is obtained at similar target/mask SF by using colour masks.

TvC curves Using a full TvC design, consistent suppression was found (Chen et al., 2000a; Mullen & Losada, 1994; Switkes et al., 1988; Vimal, 1998a). The masking effect was limited (3 times the detection threshold of luminance).

Facilitation findings were not found consistently between studies. No facilitation was found for (Switkes et al., 1988), and for most cases in Vimal (1998a), with the exception of

0.125 cpd stimuli for one participant and some other cases[21]. Chen et al. (2000a) did find some facilitation for one participant only (out of three), but this effect disappeared when using individually isoluminant stimuli instead of nominally isoluminant.[22] The results did not match the usual dipper function found in within-channel masking, when adding a luminance component to the chromatic mask, the shape of the TvC curves become more like the classical dipper-function (Chen et al., 2000a).

In general, the facilitatory effects, when found, were dependent on the participant, the SF and the contrast of mask (Vimal, 1998a). Mullen and Losada (1994) did find (suprathreshold) facilitation consistently, when both mask/target were at the same SF (with a SF bandwidth of ± 1 octave), followed by some masking (SF bandwidth of ± 2 octaves), see following sections for SF bandwidth and for local cues hypothesis for facilitation.

SF bandwidth and SF tuning Using fixed contrast and variable target/luminance SF, Vimal (1998a) did find tuned suppression somewhat similar or tighter than luminance masking. DeValois and Switkes (1983) found that the suppression was abolished at ± 3 octaves.[23]

Mullen and Losada (1994) did not find significant SF effects on the TvC shape, there was more variation between participants (only 0.25, 0.50 and 1 cpd mask/target pairs were tested). On the other hand, Vimal's (1998a) data show a peak between 2 and 4 cpd, this is confirmed by Switkes et al. (1988), which showed a clear tuning of suppression at 2-4 cpd.

Overall, this demonstrates a band pass tuning of interactions at 2 cpd, the colour masking on luminance target is more strongly tuned than the luminance-luminance (compare with Section 3.3.4). This was confirmed by testing the full TvC at 2 cpd versus 0.5 cpd, the chromatic masking of luminance is less efficient at 0.5 cpd (Switkes et al., 1988), some facilitation is visible.

Phase effect and local cues No effects of phase on iso-oriented stimulus have been found (DeValois & Switkes, 1983; Gowdy et al., 1999; Mullen & Losada, 1994; Switkes et al., 1988). This is important as it means that (equiluminant) red/green on dark/light stripes is processed similarly as on light/dark.

Using randomized phase relationships (red on dark or red on light) between trials, the facilitation effect reported by Mullen and Losada (1994) disappeared (for 2 out of 3 participants). Mullen and Losada (1994) interpreted this as the use of local cues and not facilitatory interactions between the channels. Gowdy et al. (1999) argued that this method only increases uncertainty but does not suppress the facilitatory effect. When ob-

[21] T/M: 0.5/2 cpd for two participants and 0.13/0.2 for one only, all these facilitations were at suprathreshold mask contrast.

[22] See Chapters 8 for distinction between nominally and individually isoluminant stimuli.

[23] However we have discussed previously that the contrast influences bandwidth estimation and that it requires a full TvC and a well-defined metric.

servers are asked to judge in "which interval the red bar appears brighter", the suppressed facilitation is back to normal (Gowdy et al., 1999), as the uncertainty is reduced with this explicit criterion.

Facilitation was found to disappear, by Gowdy et al. (1999), when the relative phase was 90° offset, but not by DeValois and Switkes (1983).

Pattern and edges DeValois and Switkes (1983) changed the mask to a square-wave pattern and found no effect of phase similarly to sinusoids. However, one (out of 3) participants did show reduced masking in the square-wave condition than in the phase-aligned condition. This highlights that there was phase specific processing of edges for that participant and highlights the issues of individual differences in chromato-luminance interactions (see also Mullen & Losada, 1994, for individual differences in TvC shape). Gowdy et al. (1999) generally found greater facilitation with a square-wave pattern.

Using variable pedestal contrast, Gowdy et al. (1999) did find some facilitation for sinusoidal patterns, however stronger facilitation was obtained by square-wave red-green mask (on square-wave targets) and at a low spatial frequency (0.8 cpd). Because of the randomized pedestal contrast, the observers were forced to use brightness cues and not saturation differences to perform the task.

Summary For luminance targets and chromatic masks, cross-interaction effects were found but they did not match the usual dipper-shape function. The presence of the facilitation and its cause are still debated. There is a possible tuning of interactions, highest at 2 cpd. The next section will detail responses to colour targets in the presence of luminance masks.

Luminance Mask/Colour Target

Fixed contrast mask The use of fixed mask contrast (the converse design of the one described in the above section) has yielded some limited suppressive effect (DeValois & Switkes, 1983). Vimal (1998a) did find suppression of larger amplitudes than for coloured masks. However, the suppressive strength at the same SF point was of the same efficacy as the chromatic colour mask on chromatic targets (DeValois & Switkes, 1983).

TvC function On the whole TvC function, with both mask/target of similar SF, (suprathreshold) facilitation was found followed by some masking (Chen et al., 2000a; Mullen & Losada, 1994; Switkes et al., 1988). The masking effect was limited (3 times the detection threshold of luminance). A large range of mask contrasts produce facilitation (from threshold to 30 times threshold), than masking effects at higher mask values (Chen et al., 2000a; Mullen & Losada, 1994; Switkes et al., 1988).

Once again, the cross-channel interaction is different from the classical dipper function data. The facilitation occurs at high suprathreshold contrast. By adding a chromatic component to the mask (mask is both colour and luminance defined), Chen et al. (2000a)

obtained a shift of the TvC function to the left, and reported a more classical dipper-shape form.

SF tuning Using different SF pairs for target and mask, the facilitation tuning was determined to be broad (Mullen & Losada, 1994; Switkes et al., 1988). Vimal (1998a) found a tuned SF suppression slightly tighter than within channel masking tuning, while DeValois and Switkes (1983) found that the suppressive effect was almost or completely gone at ±1 octave (using fixed contrast).

Switkes et al. (1988) found weak cross-suppression at low SF. Some facilitatory effects were found at high luminance SF (+2/3 octave DeValois & Switkes, 1983). Here the masking function seemed similar to colour/colour masking. As a main effect of SF, Gowdy et al. (1999) found that the facilitation peaked at 2 cpd and fell on either side.

Spatial cues or spatial modulation Cross-facilitation was not abolished by using randomized phase shifts (Mullen & Losada, 1994). In order to check for spatial cues (suggested by Mullen & Losada, 1994), Gowdy et al. (1999) compared unipolar (red or green bars and background bars) versus bi-polar patterns (red and green bars, i.e. standard stimuli). Those two types of patterns are detected with the same sensitivity.[24] This means that both types of bars are processed independently at that level, probably by blob detectors (see Chapter Physiology 2). However, when a luminance pedestal is present there is a two-fold improvement in detection of bi-polar versus unipolar. This would be the case if the detector were now an edge detector, taking the difference between the two bars.

Using this elegant methodology, Gowdy et al. (1999) showed that the presence of the luminance pedestal changes the colour detection from (unipolar) blob detectors, detecting only red or luminance bars/blobs, to a proper edge detector, detecting differences across an edge between red and green. This greatly increases the sensitivity. However, this should strongly link the detection to the luminance phase, and indeed it was shown that relative phases of 0° and 180° yielded equally strong effects but 90° and 270° elicited only weak facilitation (Gowdy et al., 1999).

This result is complementary to increased chromatic discrimination with pedestal using spot stimuli, as described in a following section (see Section 3.4.3). Increased wavelength discrimination has been reported to improve with the presence of a luminance pedestal (Hilz & Cavonius, 1970; Hilz, Huppmann, & Cavonius, 1974; however see also Elsner, Pokorny, & Burns, 1986). Gowdy et al. (1999) linked the increased colour sensitivity due to the luminance contrast to this improved wavelength discrimination.[25]

[24] However, green bars have a slight advantage versus red bars (Interesting parallel with Krauskopf et al., 1982; Werner et al., 2000, findings using different paradigms)

[25] Note that the alternative explanation is that the increased discrimination is due to the use of local cues, as suggested by Mullen and Losada (1994).

Pattern and edges Square-wave on square-wave facilitation was more potent than sinusoid on sinusoid, specifically for low SF for both luminance test/ colour pedestal and the opposite case. At high SF, the two patterns matched in facilitation (Gowdy et al., 1999). The facilitation was not due to the effective contrast, as tested by using higher sinusoid contrast. Therefore, it was hypothesised that the presence of the edge was responsible for this stronger facilitation. Replacing the square-wave red/green colour test by a sinusoid (with Luminance Mask) does not change the potency of facilitation, so the stronger edge is more important for the luminance pedestal.[26] However, when performing this manipulation with a luminance test and red/green mask the facilitation decreases. Gowdy et al. (1999) concluded that two types of cross-facilitation have different origins.

Linearisation of the red/green transducer Gowdy et al. (1999) performed the same experiment with varying amounts of chromatic component added to luminance pedestal, to check for chromatic artefacts. Interestingly no differences were obtained in chromatic sensitivity. They concluded that the luminance pedestal did linearise the chromatic processing. If the contrast processing is compressive (see Figure 3.10, p. 71) then adding more contrast can push the receptive units into the accelerating parts of the curve, facilitating detection, and further push into the compressive part, making detection more difficult. This is the basic principle of dipper function in pedestal experiments. However here, by adding various amounts of colour to the pedestal the detection threshold was left unchanged. Consequently, the response function (transducer) can be thought of as linear, under the influence of a luminance pedestal.

Summary It was found that cross-channel interactions with luminance mask and colour targets consistently elicited facilitation and moderate amounts of masking. As for the reverse condition, the TvC shape departs from the dipper function. The next section focuses on channel interactions as tested by spot-like stimuli.

3.4.3 Flash detection data

Cole et al. (1990) tested the detectability of a flash (1°) superimposed on another (pedestal) flash (see Figure 3.6, p. 32, on pedestals). The pedestal/mask flash intensities are varied consistently, as in masking studies with gratings.[27]

In the uncrossed condition, Cole et al. (1990) found classical dipper function results. In the crossed condition, facilitation was present for both conditions. The facilitation was mostly at suprathreshold levels and of constant strength across the range. No differences were found using positive or negative pedestals. Cross-masking was not found at higher

[26] Facilitation disappeared when the two components were made orthogonal and little to no masking was found (Gowdy et al., 1999).

[27] For metric comparison, studies with recurrent grating type stimuli are presented (Section 3.4.2) use contrast instead of raw intensity.

mask values, which is a major difference compared with grating (sinusoids or square-wave) data.

Additionally, Cole et al. (1990) demonstrated that the presence of overlapping chromatic and luminance edges could generate the facilitation. The facilitation effect can be suppressed by the absence of a surround or by using dichoptic presentation.

This suggests two types of facilitation mechanisms for crossed and uncrossed conditions. According to Eskew, Stromeyer, Picotte, and Kronauer (1991), this also suggests that there is no common detection channel with a transducer effect, Eskew et al. (1991) also concluded that the facilitation effect was not due to a reduction in uncertainty. Cole et al. (1990) concluded that the luminance pedestal linearises the chromatic response (as was found later by Gowdy et al., 1999). However, Chen et al. (2000a) suggested that the cross-facilitation for spots and sinusoids (or gabors) might have a different origin (without specifying the possible mechanism).

3.4.4 Theoretical analysis of channel interactions

In this section, I will further discuss the implications of the results described in Section 3.4.

Interactions asymmetry

Section 3.4.2 showed an asymmetry in interaction effects between channels, first reported by DeValois and Switkes (1983). The suppression appears to be SF tuned and strong across a broad range for chromatic masks on luminance targets, whereas for luminance masks/chromatic targets it appeared weaker and only at the same SF. Adaptation data does show mild differences in adapter potency, towards chromatic being a stronger adapter.

Furthermore, in terms of facilitation, Switkes et al. (1988) showed no effect for colour mask/luminance test, and a large broadband effect for luminance mask/colour test. This is where other reports are contradictory. Mullen and Losada (1994) did not find the asymmetry with similar suppressive effect as Switkes et al. (1988). It was surprising that they did report facilitation of luminance detection by colour masks, but using random phase relations, this effect disappeared[28]. Vimal (1998a) did find a "mild asymmetry" in the cross-channel masking abilities of colour and luminance as did Mullen and Losada (1994, scrambled phase experiment). Vimal (1998a) did report very little facilitation, possibly owing to the fact that the orientation between mask and target was $+14.5^{\circ}$ therefore limiting local cues, as suspected by Mullen and Losada's (1994) phase data or luminance colour edge detector (Gowdy et al., 1999). This is further emphasized by the flash detection data (see Section 3.4.3). However, Gowdy et al. (1999) did suggest an account for these phase effects (Section 3.4.2), involving two different mechanisms for cross-facilitation, producing asymmetry as well (see Gowdy et al., 1999, for more details).

[28]For 2 out of 3 participants (Mullen & Losada, 1994).

Vimal (1998a) did note that the type of interactions or even absence of interactions was dependent on test and mask SF, mask contrast with variability between observers. Vimal's (1998a) study on the slopes of cross-condition suggested weaker contrast non-linearity, which might be linked to the linearisation hypothesis (for luminance mask/colour test case only).

Mechanism Hypotheses

Several hypotheses have been formulated about cross-channel interactions. Mullen and Losada (1994) formulated three ways that interactions might be processed (see also comments on models in Discussion of Vimal, 1998a). The first model is the double duty-model, the second a model with two channels and pre-transducer interactions and a model with two channels and post-transducer interactions. Another category is a mix of models 2 & 3 (Switkes et al., 1988).

Double duty model First of all, pathway interaction is not the only conclusion possible from the results of masking/pedestal experiments (Chen et al., 2000a; Kelly, 1983). Similar effects would also occur if luminance and colour were both processed by one pathway (see also Gur & Akri, 1992; Regan, 1991). Advocates of this view note that the chromatic pathway could be detecting some of the luminance signal, through the same transducer. However, some evidence might point to the contrary. The different shape of cross-channel interaction compared to a dipper function would suggest that a common transducer processing both luminance and colour could not produce these interactions (Gowdy et al., 1999). The asymmetry in interactions reinforces this point. It is evident that within- and cross-channel facilitations are of different origin (e.g. Cole et al., 1990; Gowdy et al., 1999). Facilitation within channel is well explained by non-linearity models (see Section 3.3).

Pre-transducer interactions The second hypothesis is that the interactions occur before the transduction stage by an excitatory interaction (initially formulated by Switkes et al., 1988). Chen et al. (2000a) and Chen, Foley, and Brainard (2000b) suggested that the high contrast necessary for facilitation means that the contribution of luminance to the chromatic channel is done through a small excitatory input, thus agreeing on that point with Switkes et al. (1988). At higher contrast the low masking is also linked to small input in the divisive mechanism (small cross-inhibition).[29] Chen et al. (2000b) suggested that all channels interact by divisive inhibitory input (see Broadband inhibition in subsection on Masking above 3.3.3). The excitatory model (Switkes et al., 1988) was not found adequate, as randomized pedestal contrast did not suppress the facilitation (Gowdy et al., 1999).

[29] Chen et al. (2000b)'s model is an extension on Foley (1994)'s model for the cross-masking with Switkes et al. (1988) ideas of pre-transducer excitatory interactions.

Post-transducer interactions The third hypothesis is that the interactions happen after the transducers through a mix of excitatory and inhibitory mechanisms (Mullen & Losada, 1994).

Gowdy et al. (1999) suggested a late stage interaction between channels (see also Cole et al., 1990; Derrington, Krauskopf, & Lennie, 1984; Krauskopf et al., 1982), with a comparison between colour processed and luminance edges. The colour edges are demarcated with the help of the luminance edges, then are integrated to the luminance edges to be analysed across the edge. Demarcation and integration are susceptible to phase effects (see also Mullen & Losada, 1994; Vimal, 1998a). This process is greater with sharp edges (square-wave). The facilitation by demarcation was also suggested by Cole et al. (1990), because when the surround was eliminated facilitation disappeared.

Multiple stages interactions/Multiple mechanisms Switkes et al. (1988) additionally suggested that the low contrast and high contrast transduction could be set in different cortical regions (stages). Therefore, interactions are possible at different processing stages.[30]

Luminance facilitation on colour, and vice-versa, needs to be modelled differently as evidence provided in previous sections suggests different mechanisms at play. Furthermore, the impact of sinusoid or square-wave pattern varies between those cross-channel types; thus providing further proof of different mechanisms. The facilitation of luminance masks on colour targets is greater than the reverse case. Gowdy et al. (1999) also suggested another facilitation effect independent of edges to account for the remaining facilitatory effects.

Vimal (1998a) suggested that luminance and colour segregation at threshold is then followed by a common divisive pool at suprathreshold levels, the facilitation could be due to a suppression of inhibition (disinhibition) between channels.

Separation of channels

It seems then, that at least at the detection threshold that luminance and chromatic channels are processed in separated channels (Cole et al., 1990; Giulianini & Eskew, 1998; Stromeyer, Thabet, Chaparro, & Kronauer, 1999). Additionally, that chromatic information is processed by two independent mechanisms (or axes in colour space), red-green and yellow-blue (Krauskopf et al., 1982).

Facilitation, due to target and pedestal, being processed by the same transducer has been described previously. In short, the facilitation happens below or at threshold detection level of the mask (pedestal). This leads to the usual first part of the dipper-shape. When cross-facilitation does occur it is at suprathreshold levels (of mask detection) and over a wide range of contrast thereafter. This suggests that luminance and colour are not processed through the same transducer (Cole et al., 1990; Mullen & Losada, 1994;

[30]This model contains an accelerating non-linearity followed by a compressive one in another processing stage, essentially producing the classic transducer function.

Switkes et al., 1988; Vimal, 1998a) or in other words, there is a separation of detection mechanisms. At sub-threshold level, there is no interaction between these channels (see also additivity test, TvC with pedestal by Mullen & Losada, 1994, for further evidence of independence).

However, as stated in the above section, we need to take into consideration the possibility that some channels might be able to process both luminance and colour in some restricted set of conditions (Chen et al., 2000a; Kelly, 1983). There are some caveats, the separation of luminance and chromatic transducer might only be the case for low to medium contrast (Giulianini & Eskew, 1998; Mullen & Losada, 1994; Stromeyer et al., 1999). The suprathreshold behaviour could be due to a combination of effects. Meese and Holmes (2007) suggested at least two inhibitory mechanisms within the luminance channel: one tuned and one isotropic (with stronger amplitude at low SF). It was also suggested that two processing channels for chromatic information exist (Gheiratmand et al., 2013; Kim et al., 2013); one low-pass, non-oriented and monocular; and one bandpass oriented and binocular. Data on orientation effects between luminance & chromatic channels would be an interesting addition to the literature, as it would test possible interactions at oriented cortical sites.

Several known effects between luminance and colour interactions happen at suprathreshold (e.g. McCollough effect) and colour appearance per se might be based on multiple mechanisms (see Discussion in Giulianini & Eskew, 1998). I will detail several additional suprathreshold effects in the Section 3.5 on suprathreshold spatial processing in the context of colour & luminance interactions.

Non-cardinal Encoding

It is possible that there is some between-axes encoding at later processing stages (Krauskopf et al., 1982). Several stages of processing have been suggested (Gowdy et al., 1999; Krauskopf et al., 1982; Mullen & Losada, 1994; Switkes et al., 1988). Furthermore, as we have seen, recent data indicate two possible parallel channels within chromatic (Gheiratmand et al., 2013; Kim et al., 2013) and within luminance (Meese & Holmes, 2007) channels. There are some interactions found between the two mechanisms. Even though the two channels might be segregated at first, the observed interactions might be due to non-cardinal encoding or interactions between the mechanisms (see Clifford, Spehar, Solomon, Martin, & Zaidi, 2003).

As encoding becomes multidimensional along the processing stages, we should expect to find perceptual effects or illusions linked to specific combination of two dimensions (e.g. orientation and colour, such as in the McCollough effect described earlier). These contingent effects, where the effect only happens when two dimensions are in specific combination[31] are the results of channel interaction or integration.

[31] Contingent effect can be described as requiring a logical *and*.

3.4.5 Summary

In this section, I have listed several paradigms in which cross-channel interactions occurs. Adaptation studies showed that luminance and chromatic adaptation effects existed in combination with orientation encoding, thus suggesting the influence of a cortical site. Furthermore, masking studies showed interactions that gave rise to different TvC shape, away from the classical dipper-shape. The implications were discussed in the theoretical analysis section. The cross-channel effect sometimes appears to be asymmetrical and dependent on the type of edge. There are also suggestions that luminance linearises chromatic contrast processing and turns it from a blob detector to an edge detector (although this is debated). Most of these cross-channel interactions happen at suprathreshold and were performed in the context of detection tasks. Section 3.5 will further focus on the spatial dimension and the possible interactions with luminance and chromatic signals, this will further lead us to the last section on visual scene interpretation.

3.5 Spatial processing interactions with luminance and chromatic signals

3.5.1 Combined detection

The previous section focused on the detection of one type of stimulus in the presence of another kind. However, Gur and Akri (1992) studied combined luminance and chromatic targets. The detection task compared the detection of luminance, (red-green) chromatic or combined (when no component is a mask per se). Their results showed that the combination of the two elicited greater sensitivity than any channel alone as well as faster reaction time. Moreover, they demonstrated that the combination was greater than the theoretical combination of independent channels. This was interpreted as proof of facilitatory interactions, but can also be alternatively interpreted as contradicting independent channel theory.

3.5.2 Orientation perception, tilt illusion

The tilt aftereffect is an illusion linked to orientation encoding where the true orientation of a target is shifted in the opposite direction to a previously observed adapter. It is strongly linked to adaptation studies and it was first discovered for luminance gratings. However, we know the chromatic information is encoded spatially (see for e.g. Bradley et al., 1988, with orientation selectivity adaptation of purely isoluminant adapter/target pairs).

Flanagan, Cavanagh, and Favreau (1990) found highest tilt aftereffects when target and match were aligned in colour space, the lowest effect is when adapter/target pair was orthogonal in colour space.[32] This was interpreted as independent encoding of orientation

[32]See Chapters 4 & 8 for details on colour spaces.

mechanisms of chromatic and luminance channels. However, cross-adaptation effects were present.

Clifford et al. (2003) presented extensive chromatic data on the surround tilt illusion. In this illusion, a differentially oriented surround repulses the perception of a central target's orientation. This effect was found to work in the luminance but also the chromatic[33] domain, this once again shows the encoding of orientation by chromatic channels. As opposed to Flanagan et al. (1990), non-cardinal colour modulation (combination pairs of red-green, luminance or blue-yellow) all had similar effect potency. This demonstrates that non-cardinal axes are not special anymore at this level of processing. As pointed out by Clifford et al. (2003), the processing of surround interactions require lateral interactions in the cortex, as opposed to the reduction in feedforward input that might happen in adaptation aftereffects.

3.5.3 SF perception

If one adapts to a specific SF a subsequent test will appear shifted away from it. Favreau and Cavanagh (1981) demonstrated that chromatic and luminance SF adaptation could be independently shifted in opposite directions. They suggested independence in SF processing between the channels. However, Favreau and Cavanagh (1983) did show some cross-channel shifts, which they link to cells encoding both luminance and colour (see Physiology Chapter 2). Favreau and Cavanagh (1983) showed that this effect worked between the eyes, except when the stimuli were composed of luminance and colour.

3.5.4 Stereopsis

Random-dot case C. Lu and Fender (1972), using random-dot stereograms (Julesz, 1971), demonstrated that at equiluminance (dot purely defined by colour changes), the percept of depth was abolished in accordance with Livingstone and Hubel (1987). However, a series of papers by Kingdom, Simmons, and Rainville (1999); Simmons and Kingdom (1997, 2002), challenged this idea. They suggested that local cues (dots) were difficult to process by the chromatic channels. Indeed, larger shapes were processed adequately in stereoscopic depth.

Interestingly these two channels seem to interact in an asymmetrical manner (Simmons & Kingdom, 2002). It was possible to disrupt or improve stereo-detection using correlated or anti-correlated luminance and chromatic patterns. The purpose of the interactions between the independent stereo channels might be to unify the depth perception (Simmons & Kingdom, 2002). This interesting interpretation could be extended to contrast perception and other spatial processing.

Stereopsis in object perception Colour and stereo-information also interact in the context of reflectance extraction. The stereo-shape of an object can influence the percep-

[33]The chromatic tilt effect was not present a low contrast for some participants.

tion of the objects' colour. This was demonstrated by Bloj, Kersten, and Hurlbert (1999) using folded cards with inter-reflections. The stereoscopic depth of the object was reversed using a pseudoscope, and this change in object shape subsequently evoked a shift in colour perception of the cards. Using this paradigm, Bloj et al. (1999) demonstrated that the visual system is taking into consideration inter-reflections for the perception of surface reflectance. This is really advanced processing for the visual system to be performing. The visual system is essentially recovering effects due to the reflections and taking it into account but only because of a stereoscopic cue.

Spatial configuration effects were also demonstrated for lightness perception by Gilchrist (1977), the apparatus (see Figure 3.12) had different spatial configuration viewed binocularly or monocularly but the retinal images were the same (in term of lightness intensity distribution and relationships). There was a stark difference in perception whenever the target stimuli were perceived within the plane or within another plane, changing from light to dark percept.

In those experiments, the structure shape and the coplanarity of the stimuli informed the visual system about the probable lightness/colour of the object.

3.5.5 Summary on contingent effects

A list of contingent effects were mentioned in this chapter, I briefly summarize them here.

Firstly, colour has been found to be strongly linked to orientation (Clifford et al., 2003; Flanagan et al., 1990; McCollough, 1965; Werner, 2003; Werner et al., 2000; Zaidi et al., 1997) but also SF perception (Favreau & Cavanagh, 1981, 1983). This data overwhelmingly supports the role of colour in spatial processing of form and orientation.

Further along the processing hierarchy, the next step could be to extract texture information and reflectance information. Robinson and MacLeod (2011) showed that the McCollough effect work also with plaid stimuli; these results suggested a specific encoding of plaids additional to the grating encoding demonstrated by selective adaptive effects.

The ultimate goal of visual processing is to facilitate navigating and interacting with the environment, following this ecological constraint we can also interpret the processing in terms of a functional role. Some previously mentioned phenomena have already been linked in that way. For instance, chromatic adaptation can serve a role to discard the ambient light colour from the colour of an object. The next section will discuss scene information extraction with an emphasis on luminance and colour interactions.

3.6 Visual scene interpretation

This section is dedicated to the computational role of visual perception in terms of scene information extraction. As such it will be at a higher level on Marr' s stages of analysis (Marr, 1982) than previous sections. The computational problem of scene information extraction will be discussed. From there, I will move on to segmentation and shape-from-

shading. I will thereafter discuss how the human visual system might solve this problem efficiently using heuristics and assumptions (i.e. computational shortcuts, e.g. for gloss perception see Kerrigan & Adams, 2013).

Some studies perform scene information extraction in the luminance domain (through the luminance channel) and relegate the chromatic information to determining the colour of objects. The investment in term of colour neurons in the visual cortex (see Physiology review, Chapter 2) is such that it would make sense to use chromatic encoding for more than spectral estimation.[34] Rubin and Richards (1982) have also suggested that the chromatic information might have a role in disambiguating the changes in intensity due to reflectance (material) from those due the lightfield.

3.6.1 The computational problem

According to Marr (1982)'s levels of analysis, in order to analyse a system we must first ask what question is being solved by the computation. The ecological approach gives us a grounded approach to this question. As an interacting agent within our environment, we can reduce part of the problem to object segmentation.[35]

In order to understand a visual scene, it is required to extract visual properties. For example, Barrow and Tenenbaum (1978), formulated a computational problem as follows; they defined as input the visual array (e.g. set of intensities from the retinal projected into the cortex) and as output a set of images with characteristics. The characteristics extracted are reflectance, depth, orientation of shapes, illumination (light sources), shading (shadows through shape obstruction) and transparency (Adelson & Pentland, 1996; Barrow & Tenenbaum, 1978).[36] These were envisioned first in luminance intensity and human lightness perception. However, this could be extended to hyperspectral cases, with hyperspectral reflectance and illumination extractions.[37]

There is evidence that the visual system is performing such extraction tasks as can be seen with some illusions (e.g. Adelson, 1993; Gilchrist, 1977, for luminance, see also Logvinenko, 1999, and comments by Kingdom, 1999, as well as Lotto & Purves, 1999, for colour brightness). Human observers are able to discount shadows, understand transparency, keep colour and lightness constancy and have *shape interaction awareness*.

For example the luminance shading across an edge of an object, such as a wall or the simpler case of a cube, is not misinterpreted as a change in reflectance but understood as a differential lighting towards that objects. If asked, an observer might say that the colour is identical on both edges even though the actual *raw* inputs to the cones might be very different. An example of effect of shape arrangement and lightness perception (Gilchrist,

[34] As DeValois and Switkes (1983) put it *"Surely such an extraordinary neural investment was not made just to allow the aesthetic appreciation of sunsets!"*

[35] Of course an interacting agent requires more abilities such as pattern recognition and decision making from this processed information, these are two others main areas in psychology and artificial intelligence and are not within the scope of this thesis.

[36] This work does not explain how these cues are combined into one percept.

[37] Following Marr (1982)'s second level of analysis, the representation level, the hyperspectral colour surfaces are represented in the cortex by colour spaces (see Methods Chapter).

1977) is illustrated in Figure 3.12. Here the coplanarity that varies between monocular and binocular viewing has a direct effect on the estimate of the scene lightness.

The whole scene characteristics are taken into account whether it is through shadows (Adelson, 1993) or differential illumination (Lotto & Purves, 1999), and as we have seen with Bloj et al. (1999) the capacity to understand inter-reflection through stereopsis information about the object's shape. This particular last example is really telling as it is dependent on the 3D shape (from stereopsis), hence the different spatial characteristics that are extracted do interact with each other's.

This multiple features extraction is a complex problem to solve, the complexity comes from the fact that intensity changes can come from multiple origins (Kingdom, 2008; Rubin & Richards, 1982). The factors that can vary the intensity array are the spatial and material properties of the scene as well as the light field properties of the scene. Taken together this consists of object reflectance (with possible pattern and pigmentation changes), their shape (i.e. surface orientation changes), creating possible gradients of lights, shadows and highlights and light(-s) composition and directions (Rubin & Richards, 1982). Consequently, this is why Adelson and Pentland (1996)'s model, introduced earlier, has several outputs for each of the factors of intensity changes. These are estimated concurrently, as the visual system recovers the spatial arrangement the material properties and the light conditions (Adelson & Pentland, 1996).

The scene-recovering problem has been approached using different approximations (to relax some of the computational complexity). For example, grey world assumption (Kingdom, 2008, Figure 1); flatness (Land, 1977); shadowless (Land, 1977); uniform reflectance regions (i.e. no patterning, e.g. Kingdom, 2008; Land, 1977), one light source (Kingdom, 2008), matte surface (as opposed to one with specular highlights).

One other alternative to complexity is to restrict the complexity of the question. One of the early studies of this kind, by Rubin and Richards (1982), constrained the problem to find material borders and occlusions. Another example of the kind is the shape-from-shading, which is both a perceptual phenomenon and a restricted computational problem, this will be described in its own section later on.

Kingdom (2008) presents the problem of light versus material estimation and describes a list of heuristics that could be used by the visual system to solve this general issue. Importantly his review also takes into account texture and chromatic cues.[38] In many cases, computation or heuristics are performed at the edges (intensity changes) within a scene.

3.6.2 The role of edges in scene interpretation

Usually edges have only one origin, i.e. shadow, reflectance change, object physical edges, and co-occurrence of multiple factors at one edge is rare (Marr, 1982; Marr & Hildreth, 1980). That knowledge can be used as a powerful shortcut. Therefore, edges have a

[38] See also Ramachandran (1991, "bag of tricks" comment).

specific significance and can be used as a source of information about the visual scene (Land, 1977).

The distinction between smooth changes and abrupt ones can potentially be used for segmentation, as sharp edges are often associated with reflectance changes while illuminations changes tend to be smooth (Horn, 1974). Taking the intensity ratio between two spatially close points is the simplest way to perform this computational task and can be used to detect an edge (Land, 1977), by having three receptors (trichromacy) at this two spatial point Rubin and Richards (1982) determined that this was sufficient to distinguish between reflectance and other types of changes.[39] Land (1977) suggested that taking this (spatial) ratio consistently across the visual field would result in an image of the edges that discounts non-uniform illumination and results in ratios of reflectances. This however does not take into account the possibility of sharp changes in illumination (such as shadows). Subsequent research has shown that the visual system uses slightly more complex spatial functions (see Chapter 2) to extract edges, e.g. double-opponent cells that can be modelled by gabors or DOG.[40]

In Section 3.5.4 it was shown that spatial arrangement (3D structure) could influence lightness and chromatic perception. Gilchrist, Delman, and Jacobsen (1983) reformulated the computational issue focusing on edges, namely how human observers distinguished between reflectance changes and illuminations changes (when an edge is observed). In Gilchrist et al. (1983), similar retinal inputs were created using two conditions, one reflectance only and the other reflectance plus illumination. The targets were perceived differently because the edge attribution was different. Consequently, local computation on its own cannot be the main driver of light perception without proper scene interpretation (see also Land, 1977). Once again, it appears that the visual system works out the illumination and accounts for it, but the *raw retinal values* are not perceived directly.

We have seen that two things are required: scene awareness and some knowledge of light interactions. Sharpness of the edges might play a role in edges classification (Horn, 1974).

Edge type: sinusoids versus square-waves It has been suggested that the type of edges encountered might be linked to the cause of that edge (illumination or reflectance). From the psychophysical data we have seen that square-waves tended to increase interaction effects (e.g. Gowdy et al., 1999). Therefore, it is conceivable that the visual system is treating sharp edges and sinusoids differently to extract different type of information. I will explore this hypothesis directly in my first experiment (see Chapter 5).

3.6.3 Suprathreshold contrast perception

The processing of both luminance and chromatic signals at suprathreshold levels has been shown to be different from threshold level perception (Vimal, 2000). It has been suggested

[39] Tetrachromacy only moderately increases performances (Rubin & Richards, 1982).
[40] See examples of DOG as spatial filters in Figure 2.2, p. 9.

that contrast perception and colour appearance might be complex (Cannon & Fullenkamp, 1993) and performed by different neural populations (Switkes & Crognale, 1999).

Apparent contrast of colour and luminance stimuli Using a contrast matching paradigm,[41] luminance and chromatic (red-green) channels were found to be band pass at low contrast i.e. at equal contrast a sine-wave of 2 cpd will be perceived to have more contrast than a 0.5 or a 5 cpd stimulus.[42] However, this effect disappears at high contrast and the matching contrast functions becomes flat for both luminance (Georgeson & Sullivan, 1975; Tiippana & Näsänen, 1999) and colour (Tiippana, Rovamo, Näsänen, Whitaker, & Mäkelä, 2000; Vimal, 2000) channels. Importantly the CSF and CMF (*contrast matching function*) might be independent of each other (Georgeson & Sullivan, 1975).

Switkes and Crognale (1999) demonstrated that observers were able to match contrast amplitude between luminance and chromatic stimuli despite their difference in appearance and early processing separation. Red-green stimuli were found to have a stronger perceived contrast.

Centre-surround effect The perceived contrast of a central target is dependent on the surround (J. A. Solomon, Sperling, & Chubb, 1993), with usually a lower perceived contrast in the presence of a surround (doughnut) as opposed to uniform background. Iso-oriented effects are usually suppressive but can be enhancing in some cases (Cannon & Fullenkamp, 1993; Ejima & Takahashi, 1985; J. Xing & Heeger, 2000).

Cross-modulation effects are present too, but tend to be weaker than iso-oriented effects (Cannon & Fullenkamp, 1991a; J. A. Solomon et al., 1993; Yu et al., 2001) and often facilitatory (J. Xing & Heeger, 2000; Yu et al., 2001). Yu et al. (2001) reported large individual differences in direction and strength of XOM effects. In the periphery suppression gets stronger and facilitation effects weaker (J. Xing & Heeger, 2000), with large individual differences.

Using low SF stimuli, Meese and Hess (2004) did find consistent suppression with annulus (for binocular, monocular and dichoptic presentations).

To reconcile data from Cannon and Fullenkamp (1991a) and Ejima and Takahashi (1985), Cannon and Fullenkamp (1993) performed contrast matching on a greater number of participants[43] and observed "striking individual differences", from suppression to strong facilitation. This was not due to variation in detection threshold. They suggested the need for more research on the characteristics of the excitatory systems (in term of orientation and SF).

Some of these effects might be due to brightness induction Ejima and Takahashi (1985), by using a gap between centre and surround it is possible to reduce induction effects (Yu

[41] A stimulus target is presented next to an adjustable match (Georgeson & Sullivan, 1975) or the observer indicates the highest contrast in a 2IFC design (Vimal, 2000).

[42] The contrast falls more steeply at high SF (Georgeson & Sullivan, 1975).

[43] n=10.

et al., 2001).

Cannon and Fullenkamp (1991b) used a modified version of Legge and Foley (1980)'s model to account for suprathreshold contrast perception, linking it to interactions between transducer functions (Wilson, 1980). The suppressive component of the model is dependent on pooling among filters. Their model predicted orthogonal suppression at low level but not at "high" contrast values (>1 %).

Superimposed masking Meese and Hess (2004) were the first to measure suprathreshold contrast perception of one component in a masking paradigm. Suppression was present for dichoptic and monocular, however, many cases are facilitatory in the binocular presentation (Meese & Hess, 2004, Figure 5 & 6).

Contrast Induction Singer and D'Zmura (1994) studied temporal contrast induction of surround on perceived centre contrast. Increased induction was found with increasing surround size (spatial pooling) saturating at 4°, no effect of relative orientation. Interestingly, achromatic surround did modulate the contrast of chromatic stimuli (red-green and yellow-blue) but not the opposite (i.e. interaction asymmetry). Chromatic channels did modulate each other. However, the highest contrast induction was always within the same channel. Singer and D'Zmura (1994) took it as a proof of non-independent contrast gain control between channels.[44] Singer and D'Zmura (1994) found no orientation selectivity at high contrast modulation.

3.6.4 Contour extraction

Contour integration can be performed by both luminance and red-green channels (McIlhagga & Mullen, 1996; Mullen, Beaudot, & McIlhagga, 2000) and more efficiently by luminance Mullen and Beaudot (2002). This integration to detect shape could also be performed with alternating gabors but leads to performance loss (McIlhagga & Mullen, 1996). Chromatic shape integration can nevertheless reach hyperacuity levels (Mullen & Beaudot, 2002) and the differences in performance between luminance and chromatic stimuli were linked to the poorer orientation encoding of the chromatic channel.

Additionally, McIlhagga and Mullen (1996) tested the ecological argument of combining luminance and chromatic information for object segregation. When gabors were composed of combined luminance and chromatic contrast, the ecological hypothesis tested did not work and considerable individual differences were found.

Mullen et al. (2000) later argued for a common mechanism to contour extraction, however shape analysis might be different between luminance and chromatic mechanisms (Mullen & Beaudot, 2002). The computational problem raised by object identification is the binding of features into an object (Mullen & Beaudot, 2002). Segmentation plays an

[44] Singer and D'Zmura (1994) still included the possibility of in-between axis encoding.

important role in extracting information and it can be obtained in several ways. Once the contours are extracted further processing can be used to inform the scene interpretation.

3.6.5 Effect of shape on lightness perception

We have seen some examples defined previously where the stereo-defined shape influences the perception of colour. Knill and Kersten (1991) demonstrated a shift of lightness perception when a contour defining a shape is added to a base stimulus. In essence, this is also similar to Gilchrist et al. (1983), when the retinal input was similar but the lightness varied. These two examples are with purely *stereo-flat* stimuli.

This demonstrates that the perception of shape contour *explains* the variation of intensity on the surface, and so, such variations are discarded as shading. In order to recover the actual reflectance of the object, the visual system discards variations due to shape.

The next section will define a specific scene extraction process, where the luminance variations are interpreted as variations of shape.

3.6.6 Shape-from-shading

Introduction

Natural objects of a uniform colour, when imaged, can show gradients of lights across their surface. This is due to the changes in surface orientation with respect to the light source. Hence, in principle their luminance gradients can tell us about object shape, assuming a constant known (or inferred) light source.

Given an object of uniform reflectance it has been shown that variations in luminance can be used to obtain information about object shape, this was originally formulated by Horn (1975); Pentland (1982a, 1982b) and see also Horn and Brooks (1989). The extraction of shape information purely from shading is called shape-from-shading (*SFS*).

Three-dimensional shapes inferred from 2D images can be perceived as ambiguous (Curran & Johnston, 1996). This is perhaps not surprising as luminance and chromatic gradients in a scene can be due to a host of factors: material or reflectance changes (e.g. the patterning on the surface of a leaf), differential reflection of light due to object shape (which we can interpret as shape-from-shading), shadows and inter-reflections. The variations of intensities in the retinae can be due to a host of reasons, gradient information is therefore inherently ambiguous. Information from other sources might therefore be helpful in disambiguating what luminance delivers so that shape can be extracted (see recent review by Kingdom, 2008).

Nevertheless, human observers are able to use this cue to estimate shape (e.g. Todd & Mingolla, 1983, showed that observers can extract information about curvature from shaded cylinders), see Schofield, Sun, and Mazzilli (2013) for review on human SFS. Perceived shape can be affected by changes in the scene that are not related to the object for example, by varying the lighting direction (Nefs, Koenderink, & Kappers, 2005). However, under certain circumstances, luminance and/or chromatic gradients can act as a

robust cue to shape, e.g. they can be optimally combined with each other (Harding et al., in prep), with binocular disparity to improve shape perception (Lovell, Bloj, & Harris, 2012). This enables the relationship between light direction, object reflectance and shape to be quickly learned (Harding, Harris, & Bloj, 2012). Gradient information can be combined with chromatic information efficiently. Harding et al. (2012) showed that an observer could quickly learn the relationship between light direction, the object reflectance and the shape of an object. Gradient information can also be combined with binocular disparities cues. Lovell et al. (2012) showed that this can be done optimally and fast (Lovell et al., submitted). Observers not only use the luminance gradients but also more complex luminance information such as second order luminance variations (Schofield, Hesse, Rock, & Georgeson, 2006).

Most early models of shape from shading, as well as experimentation with human observers, were done using monochromatic stimuli and uniform reflectance. These assumptions will be explored in more detail in the following two sections. I will also detail assumptions regarding the shape.

Visual system assumptions

There is an early suggestion, in the literature on SFS (shape-from-shading) that the visual system is using assumptions to interpret luminance variations (e.g. Ramachandran, 1988). In this section, I will review some of these assumptions regarding the light field and the usual shape of objects.

Light field assumptions Ramachandran (1988) suggested that the visual system estimates only a global illuminant (point) source and also recognized the role of occluding edges as a strong cue to shape. Tyler (1998) suggested that the visual system is actually assuming diffuse lighting conditions, unless explicit information about light source is available.

J. Sun and Perona (1998) showed that observers have preferential light direction assumptions to compute SFS (always above), and the specific angle was linked to handedness and expectation of light in the visual world. Langer and Bülthoff (2000) showed that the visual system might prefer light from above versus light from below. Using their task, concave versus convex, performance with lighting from above was as good as with diffuse lighting. Consequently, the visual system must know what kind of lighting condition it is facing and not just *assume* above lighting. Observers used a better solution than *dark-means-deep*[45] to judge depth along a shaded objects surface, which means the SFS algorithm must use more complex cues.

Different type of illumination will give different luminance profiles and a one-solution-fits-all heuristic is difficult to suggest. Consequently, to extract shape-from-shading it is required to extract the illumination information at the same time to have an exact

[45] *Dark-means-deep* is a simple heuristic that transforms direct lightness data into depth data, dark being further from the observer and bright closer.

reconstruction. This involves different modules that could work in parallel (e.g. Barrow & Tenenbaum, 1978).

Shape assumptions The visual system also seems to have a prior sensitivity for convexity as opposed to concavity perception (Kleffner & Ramachandran, 1992; Langer & Bülthoff, 2001). The hollow mask illusion (von Helmholtz, 1867) is an extreme case for this prior applied to face perception. Even though shading and stereopsis[46] information indicates concavity, the prior for normal (i.e. convex) faces overwrites the perception. Observers perceive the mask as convex. The assumption of light from above and convexity might be linked together as statistically in natural scenes convex objects tend to be lit at the top and darker at the bottom (Potetz & Lee, 2003) and vice versa for concave objects. Similar asymmetrical processing between light from above and below has also be shown with Macaque (M. A. Smith, Kelly, & Lee, 2007), which might suggest that they possess similar priors.

Perception of shape from luminance-defined sinusoids

Tyler (1998) worked out the different luminance profile obtained from a sinusoidal (lambertian) object against different types of illuminations (point source at two angles and diffuse illumination). With a point-source from above the stimulus, the fundamental frequency of the stimulus is doubled on the luminance pattern, as shown in photographic pictures of Lambertian gratings (M. J. Wright & Ledgeway, 2004).

Inversely sinusoidal luminance variation will be interpreted as corrugated depth profile, even though it is not matching the actual luminance profile of a corrugated object (Tyler, 1998). Its depth perception appears also rounder than a sinusoidal function (see comment on matching stimulus in Kingdom, Rangwala, & Hammamji, 2005).

The colour-shading effect

Kingdom (2003) showed that a chromatic grating (red-green) could modulate the perception of SFS produced by a luminance grating. The experiment measured the depth perceived in an oblique luminance component through SFS and looked into the effect of the spatial properties and contrast of additional chromatic components. It specifically looked into alignment of variations in colour and luminance contrast and the reported depth perception. The alignment could be orthogonal or iso-oriented. The contrast of the luminance component was varied as well as the contrast of a chromatic component either orthogonal or aligned with that luminance component. Additionally, a second chromatic component of fixed contrast was present. Its orientation relative to the luminance component depended on the condition but it was always oblique to the other chromatic component.

[46]Note that the illusion works really well without stereo-information with purely shading cues. The visual system has fewer cues to downweight.

The effect of chromatic contrast (on the variable chromatic component) was shown to be enhancing the depth perception when it was orthogonal to the luminance component and suppressing it when the two were aligned. Importantly for that experiment, Kingdom (2003) was always used three components, one luminance and two chromatic. One chromatic component (or grating) was always aligned with the luminance grating and the other chromatic component was orthogonal to it. This results in the two chromatic components always being orthogonal to each other (plaid pattern). Depending on the condition, one chromatic gratings contrast was fixed and the other's contrast was adjusted to see its effect on depth perception.

Kingdom (2003) concluded that the orthogonal colour luminance component increased depth perception while aligned component suppressed depth, which was linked to heuristics for the extraction of reflectance and shape. According to Kingdom (2003), this effect reflects the visual system trying to explain the occurrence of luminance variation. If luminance variation co-occurs with colour then we can assume reflectance change and thus suppress SFS. However, when the two are non-aligned the visual system is thought to reinforce SFS, so depth perception is enhanced.

However, the case for the enhancement data was made with essentially two orthogonal luminance components and one colour component aligned in the direction non-judged for depth. As a general statement, the three components design is difficult to interpret due to possible interactions between orthogonal within channel effects. A two component experiment was performed in Kingdom et al. (2005), for orthogonal only (Experiment 1), but interestingly Kingdom et al. (2005) tested both L-M (red-green) and S (yellow-blue) cone modulation and found similar results, i.e. enhancement. This enhancement was maximal at 15% luminance contrast. In experiment 2, they tested the suppressing effect by adding an additional component. The results demonstrated that L-M and S stimuli had similar suppressing effects and could suppress depth enhanced by the opposite colour stimuli. The first orthogonal colour component enhances the depth of the luminance, the second colour component of a different type and aligned to the luminance suppresses it.

Kingdom, Wong, Yoonessi, and Malkoc (2006) showed that, colour contrast, luminance shading and additional luminance texture interacted with each other in the percept of corrugation depth. In this paper, the colour component was always in the same orientation as the luminance variation, the texture cues[47] produced "textures bars" oriented in the orthogonal direction. There is a phase dependency of the luminance contrast (Kingdom, 2003; Kingdom et al., 2005, see also), however with texture only (no shading) there is no phase dependency effect.

3.6.7 Neural correlates of shape from shading

Hou, Pettet, Vildavski, and Norcia (2006) looked at the EEG recordings of observers

[47]The texture was produced by using an array of gabors and modulating their orientations in function of their position to create the precept of a corrugated object, see Kingdom et al. (2005), Figure 1 in Colour Plate VI.

watching stimuli transitioning between 2D and 3D interpretations (described as bi-stable[48]). The author found a stronger positive peak at 100ms (P100) for 3D shape and a lasting negativity between 200-400 ms which they hypothesized to be a kind of visual memory. This result suggests a relatively fast processing of cue combination, cues in this instance being the outline and the shading. This finding was supported by single-cell recording in the macaque, M. A. Smith et al. (2007) showed early differential processing in shape-from-shading pop-out effects. The authors found that a third of neurons had their activity modulated by the type of SFS stimuli outside their receptive field, when the stimulus within the receptive field is different from the rest. This pop-out modulation is mostly facilitatory. Behaviourally, the macaques were faster to detect an odd light-from-above sphere in a field of light-from-below spheres (equivalent to concave sphere), than the opposite (M. A. Smith et al., 2007).

3.7 Summary

In this introduction, I have shown how the two channels processing luminance and chromatic (red-green) information are processed in initially separated channels, which interact at higher processing levels. This was shown through detection paradigms, adaptive effects and also spatial processing of information, such as orientation and SF. I then described how these two channels might be involved in the extraction of edges, the processing of suprathreshold contrast as well as the perception of colour and lightness through adaptive mechanisms. All these phenomena link to the second half of this chapter.

I have then tried to demonstrate how the visual array can be used to extract information such as shape illumination, reflectance and object layouts, related to the visual scene. Additionally, considering the complexity of the computational problem, I have detailed some of the shortcuts found to process information efficiently to extract for example shape or reflectance and introduced the colour-shading effect, a perceptual phenomenon that has been hypothesized to be linked to such computational shortcuts or heuristics to process reflectance and shape.

[48]These two interpretations were created using a congruent shading and border/outline (for shape) and an incongruent condition. The shading was 95% contrast luminance square-wave and the border was triangular jagged type object.

(a) Left: Interval without target. Right: interval with target

(b) Left: Interval mask only. Right: Mask & Target. Mask contrast is twice target contrast.

(c) Left: Interval mask only. Right: Mask & Target. Mask contrast is twice target contrast. Mask is oriented +14.5°.

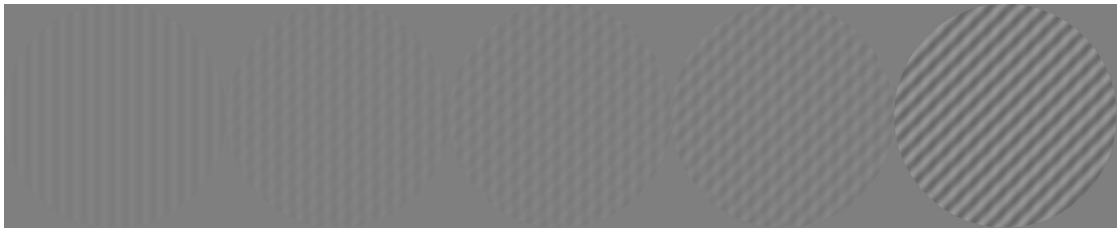
(d) Left: Interval mask only. Right: Mask & Target. Mask contrast is quadruple target contrast.

(e) Left: Interval mask only. Right: Mask & Target. Mask contrast is quadruple target contrast. Mask is oriented +14.5°.

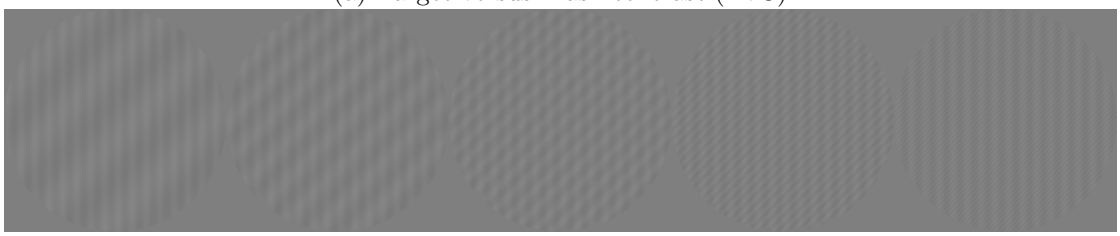
(f) Left: Interval mask only. Right: Mask & Target. Mask contrast is eight times target contrast.

(g) Left: Interval mask only. Right: Mask & Target. Mask contrast is eight times target contrast. Mask is oriented +14.5°.

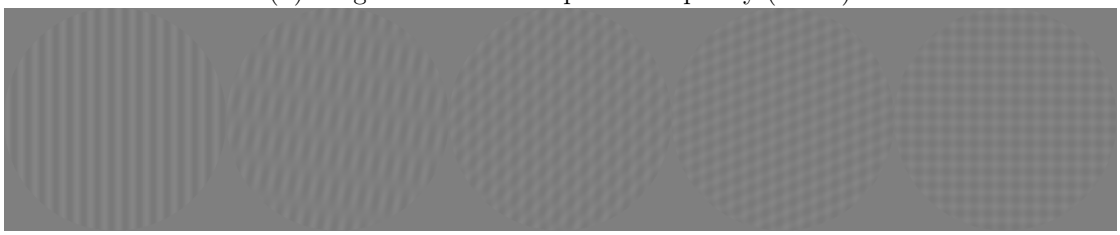
Figure 3.7: Example of iso-oriented target and mask. The detection of a target on a background (or mask) will be altered by the mask properties (e.g., contrast and orientation). The target is always present on the right side of each panel. The left side of each panel shows the mask-only condition and the right the mask plus target. The first row is equivalent to a simple detection task with no mask; other rows show detection tasks on a background. The left column shows a pure contrast discrimination task whereas the right column with non iso-oriented mask introduces spatial discrimination too. Detection of the targets becomes more difficult as mask contrast increases.



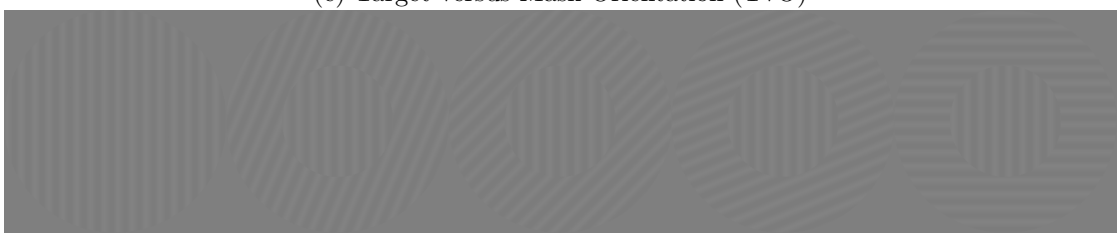
(a) Target versus Mask contrast (TvC)



(b) Target versus Mask Spatial Frequency (TvSF)



(c) Target versus Mask Orientation (TvO)



(d) Target versus Surround Orientation

Figure 3.8: Example of mask/target configurations in masking experiments. Here, for illustration, the target is always vertical and of similar contrast. In actual experiments, the contrast of the target will be varied to estimate detection threshold (sensitivity). This figure illustrates the main different types of systematic modulations applied to the mask to test its effect on the detectability of the (vertical) target.

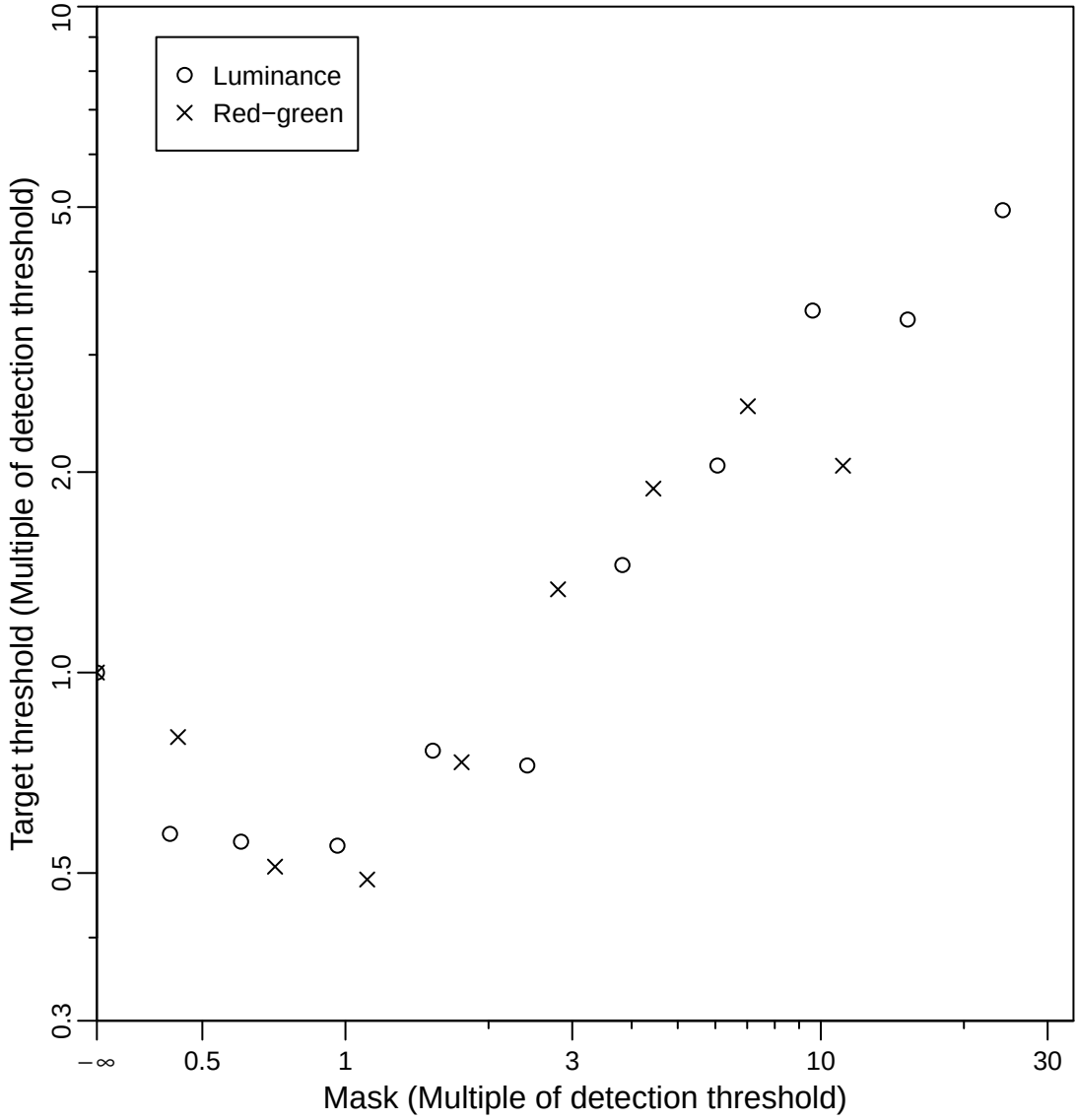


Figure 3.9: Example of Dipper function obtained for luminance and colour, target/mask pairs (Data from Supplemental Material, Chen et al., 2000a).

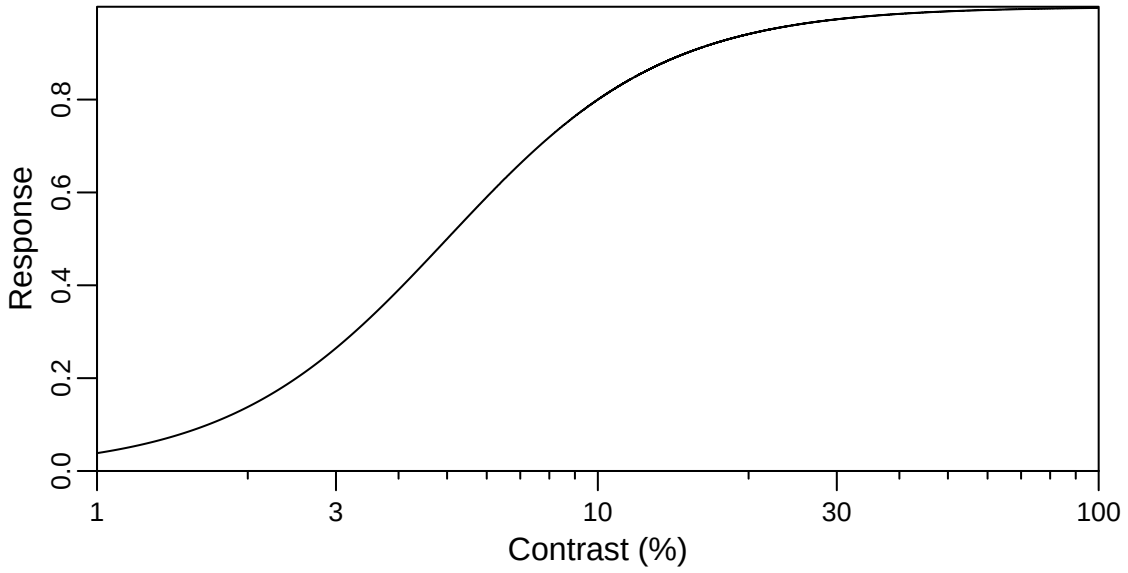
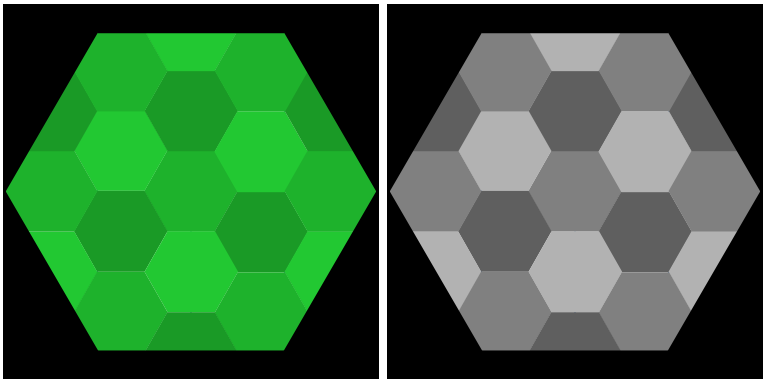
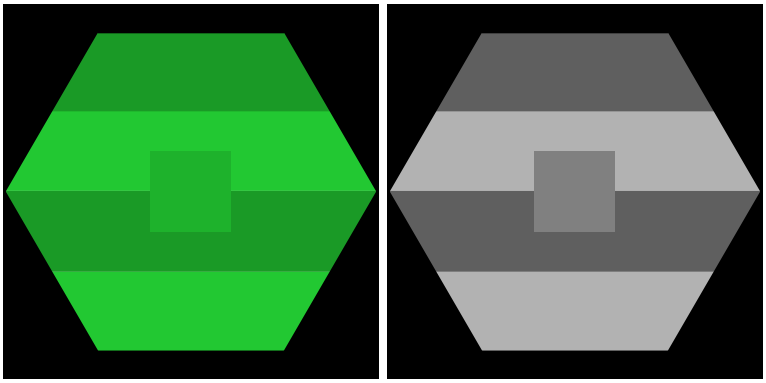


Figure 3.10: Illustration of the compressive nature of contrast encoding. Data obtained using Naka-Rushton model and parameters from Ross and Speed (1991).



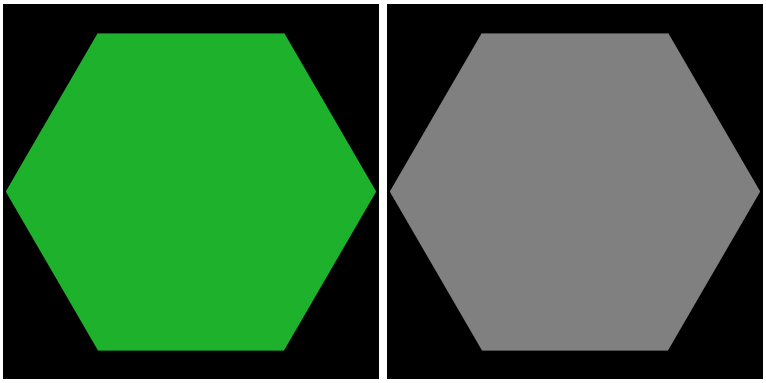
(a) Hexagon pattern, monochromatic

(b) Hexagon pattern, achromatic



(c) Stripe pattern, monochromatic

(d) Stripe pattern, achromatic



(e) Uniform pattern, colour

(f) Uniform pattern, grey

Figure 3.11: Illustration of stimuli used in Werner (2003) and Werner et al. (2000). Adaptation to stimuli on the left column will produce a shift in perception towards an achromatic percept (right column). The spatial structure (increasing in complexity from bottom to top row) accentuates the adaptation effect (Werner, 2003; Werner et al., 2000).

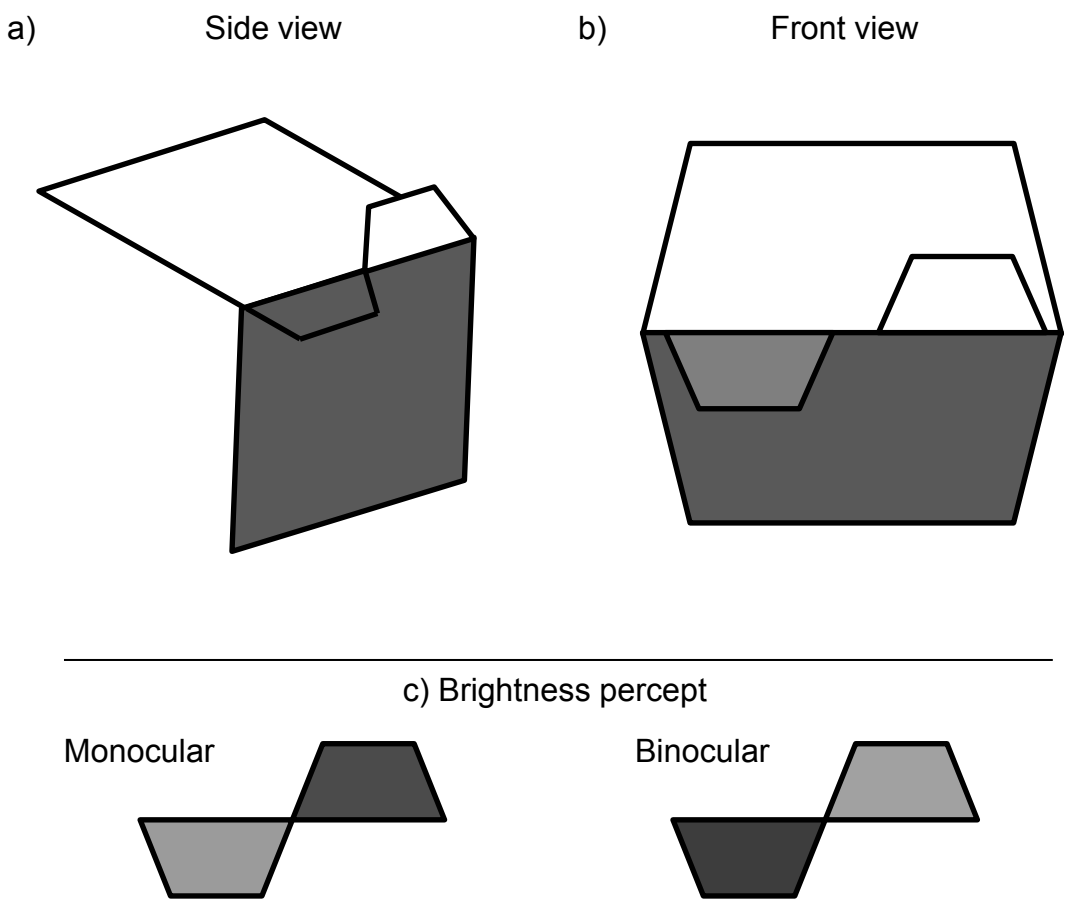


Figure 3.12: Illustration of Gilchrist (1977). a-b: views of the setup. c: The lightness percept changes depending on monocular or binocular view, as the perceived coplanarity of the targets changes.

Chapter 4

General Methods

4.1 Introduction

The aim of this section is to present the methodology and techniques used to design the experiments in the thesis. This present chapter is divided into two main parts, one theoretical introduction to standards and metric spaces used in this thesis and a second part focusing on the method used to create and display the stimuli based on these metrics.

The first half is a short primer in photometry, techniques related to the perception and measurements of brightness, as well as another ensemble of techniques and standards focused on colour perception (colorimetry). These two fields of study differ from radiometry, which focuses on the physical description of light independently of human perception. As I will show, photometry and colorimetry are heavily based on results from psychophysical experiments that defined standard observers. These standards are extremely useful in the study of human colour vision and were a major source of advancement in the field. The standard metrics used are based on historical findings and experiments that have been improved incrementally. These will be detailed. This section also introduces contrast metrics.

The second half of this chapter details more practically the method used to create the stimuli. Most stimuli used in this thesis were compound stimuli composed of oriented sinusoids or square-wave modulated in luminance or chromaticity. This section will detail the stimuli from their mathematical definition in cone contrast space to the display on a calibrated screen.

4.2 Photometry and colorimetry background

I first start by describing some basic physics of light and light emission.

4.2.1 Light emission

Any physical body with a temperature T (in Kelvin, K), emits electromagnetic radiation. The emission spectrum is directly linked to the temperature T of the object. A blackbody

is an idealized physical body that absorbs all radiations equally. According to Planck (1914), the total amount of emission M_e(in Watts per cubic meter, $W.m^{-3}$) of a blackbody at a wavelength λ and at a temperature T (K) is defined by:

$$M_e(\lambda, T) = \frac{2\pi hc^2}{\lambda^5}(\exp^{(hc/kT\lambda)} - 1)^{-1} \tag{4.1}$$

h is the Planck constant (joules per seconds, J/s), c is the speed of light (in meters per second, $m.s^{-1}$) and k is the Boltzmann constant ($J.K^{-1}$).

Previous theory thought that radiation intensity was directly proportional to wavelength $B(\lambda, T) = 2ckT/(\lambda^4)$. However, this would predict infinite energy at extremely short wavelengths (λ) and did not match empirical data. This specific distribution, described by Equation 4.1, is due to the quantum property of light, where only certain energy levels can be emitted.

The quantum nature of light was proved (latter on) by Einstein (1905) working on the photovoltaic effect. The energy of a photon is linked to its frequency $E = hc/\lambda$ and the Plank constant. Planck' s constant defines the smallest quantum, the other quanta are just multiples (as can be seen from the Blackbody Equation 4.1)[1]. The spectral distribution emission of a light source can be related to the idealized blackbody emission; a specific hue is said to have a colour temperature of x (K) if the hue is a match to the spectral distribution of a black body radiator at $T = x$ (K). Figure 4.1 shows an example of the emission at 6500K (dotted line).

4.2.2 Units and colour space

In vision science, the core methodological issue is to present a set of stimuli precisely in order to enable methodological reproduction, which should yield similar results. In the plain interpretation of psychophysics, the aim is to match psychological percepts to physical units. In the following section, I describe some of the common units and spaces.

4.2.3 Physical light description

Any light can be described by its physical value, the radiance (Table 4.1 gives a summary of the units used in radiometry and photometry, some of which will be described in this section and the following one). The radiance is the amount of electromagnetic radiation ($W.sr^{-1}.m^{-2}$) at different wavelengths.

The relative spectral distribution (consequently unitless) of daylight as defined by the CIE (Commission International de l´Éclairage) is shown in Figure 4.1. These measurements were compiled by Judd, MacAdam, and Wyszecki (1964) based on 622 samples of daylight. This has been found to be equivalent to a 6500K black body radiation (hence the name D65); however, the distribution is not as smooth as an ideal black body emitter is.

[1]Both Plank and Einstein received Nobel prices in physics for their contribution described above, which became the basis of quantum physics.

Table 4.1: Standard Units used in Radiometry and Photometry

Name	Symbol	Description
Watt	W	Standard unit of power ($J.s^{-1}$)
Steradian	sr	Solid angle, define the proportion of a sphere's surface it is defined by $\omega = S/r^2$, S = portion of sphere surface (m^2), r=radius(m). Consequently 1 sr is the surface defined by a cone section of area r^2
Radiant power	P_e	Power of electromagnetic radiation (W)
Radiant intensity	I_e	Power of electromagnetic radiation per solid angle ($W.sr^{-1}$)
Radiance	L_e	Power of electromagnetic radiation per solid angle per projected area ($W.m^{-2}.sr^{-1}$)
Luminous power	F_v	Perceived power of light (i.e. visible electromagnetic radiation) in lumen (lm). Calculated from the Radiant power (P_e) and the luminous efficiency function $V(\lambda)$ (see Figure 4.2)
Luminous intensity	I_v	Perceived power of light per solid angle in candela (cd) by definition equivalent to $lm.sr^{-1}$
Luminance	L_v	Perceived luminous intensity per projected area ($cd.m^{-2}$) or Perceived power of light per solid angle per projected area ($lm.m^{-2}.sr^{-1}$). Full calculation given by Equation 4.2

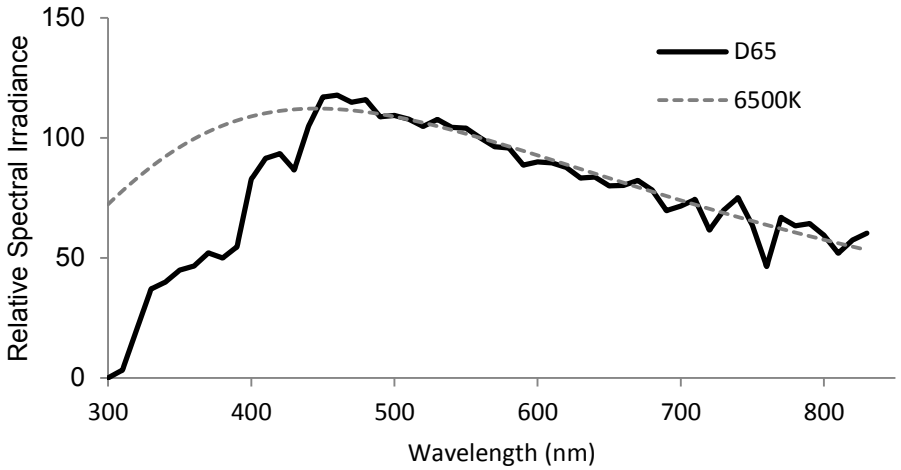


Figure 4.1: Relative spectral irradiance of illuminant D65. This represents the relative spectral distribution of typical daylight as defined by Judd et al. (1964) relative to 560nm (i.e. energy at $560\,nm = 100$) limited to a range of frequency [300-850] close to the visible range. This has been correlated as a colour temperature of 6500 K. A blackbody emittance spectrum at 6500K calculated using Equation 4.1 and normalized at 560 nm) is presented (dotted line) for comparison.

By using a spectroradiometer (such as the PR-650 used for the calibration of this thesis's experiments), it is possible to obtain these measurements for any light source. Note that the sampling of the wavelength will have an effect on the overall amplitude (see Brainard, 1997; Pelli, 1997), as a direct effect of binning. This measurement is not dependent on the observer and is very useful to fully describe a stimulus and calibrate monitors. However, it is known that the visual system samples the electromagnetic spectrum only in a narrow range (the visible spectrum, $380 - 750nm$) and the perception of human visual system is not constant within this range.

4.2.4 Describing luminance

One common unit used in vision research is luminance. It represents the perceived brightness[2] of a visual stimulus. It is necessary consider this as the perceive brightness of a monochromatic light varies with the wavelength. The photopic luminous efficiency function (see Figure 4.2) is the ratio of radiant flux between a monochromatic light λ and the light of maximum efficiency λ_m (Wyszecki & Stiles, 1982). The luminous efficiency function is therefore 1 when $\lambda = \lambda_m$. For daylight vision (referred to as photopic conditions), light is transduced via the cones and the maximum luminance efficiency is at 555nm (hence $\lambda_m = 555$). For dim light condition (referred to as scotopic conditions), light is transduced via the rods and the maximum luminance efficiency is at 507nm (hence $\lambda_m = 507$). Photopic ($V(\lambda)$), and scotopic ($V'(\lambda)$) functions, are presented in Figure 4.2.

The difference between the scotopic and photopic luminance functions has some direct perceptual effects. For example, at low light levels blue colour will appear brighter than at higher ambient light levels (described as the Purkinje Shift). The photopic standard has been updated over the years by the CIE, as can be seen in Figure 4.2. Luminance values, defined in candelas per metre squared ($cd.m^{-2}$ equivalent to $lm.m^{-2}.sr^{-1}$, see Table 4.1), can be obtained using the following equation (see Wyszecki & Stiles, 1982, pp 157 and 259, for more details on photometry):

$$L_v = K_m \int_\lambda L_{e,\lambda} V(\lambda) d\lambda \tag{4.2}$$

where $L_{e,\lambda}$ is the spectral concentration of radiance (in $W.sr^{-1}.m^{-2}$), $V(\lambda)$ the (photopic) luminous efficiency function (see Figure 4.2) and $K_m = 683lm.W^{-1}$. The luminance L is defined in term of units in $lm.sr^{-1}.m^{-2}$ more generally we use $cd.m^{-2}$, as by definition $cd = lm.sr^{-1}$. Table 4.1 provides a summary of units used in radiometry and photometry.

4.2.5 Describing colour

The fact that humans possess three types of cones as described in the Introduction (Chapter 2) has implications for colour perception. One of these implications is that any colour

[2]The term brightness tends to be avoided in vision science as it has its own meaning (Wyszecki & Stiles, 1982). Brightness is defined as a unitless measure of intensity of a test relative to a reference.

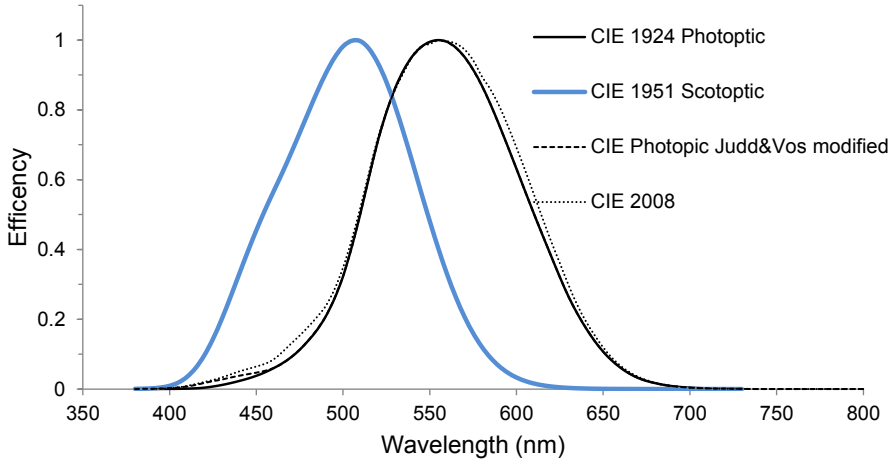


Figure 4.2: Luminous efficiency functions for Photopic (bright light) and Scotopic (dim light) vision as defined by CIE standards: 1924, 1951, and updated versions CIE Judd Vos modified and CIE 2008. A luminance value ($cd.m^{-2}$) can be obtained for any light by integrating that light spectrum with the photoptic function displayed here, see Equation 4.2

can be matched by a mixture of three monochromatic colours, often called primaries. The relationship between the mix of primaries and achievable colours has been studied and linked to human trichromacy. This has long been recognized with work dating back from the 19^{th} century by Young (1802) and von Helmholtz (1867).

Young (1802) recognized that it was impossible to have a receptor in the retina for each colour an observer can perceive at each point on the retina and suggested a much simpler and elegant solution, the three receptors solution. The selection of primaries and its implication in term of achievable colour mixture have since been studied extensively in colour science (for a more detailed history of colour science see for example Mollon, 2003, and for a detailed review on colorimetry primary selection see Wyszecki & Stiles, 1982). Using three lights (red, green and blue), Guild (1931) and W. D. Wright (1928), in separate experiments, determined the amount of primary lights necessary to match monochromatic lights in the visible spectrum (see Figures 4.3a and 4.3b).

In these experiments, an observer would be presented with a circular 2° patch split in two: one half was the target and the other half an adjustable match. The aim was to make the two halves perceptually equivalent in term of colour. The targets are always monochromatic lights, and as such, they are defined in term of their wavelength $C(\lambda)$. The matching half is composed of a mixture of three monochromatic lights usually red, green and blue (R=650 nm, G=530 nm & B=460 nm for W. D. Wright, 1928).

The mathematical representation of a match is: $C(\lambda) = r(\lambda)R + g(\lambda)G + b(\lambda)B$. The three amounts of primary lights used to make the match as a function of the monochromatic matching light $C(\lambda)$(for each wavelength λ) are called the chromaticity coordinates $(r(\lambda), g(\lambda) and\, b(\lambda))$. The chromaticity coordinates are shown in Figure 4.3, note that the red function goes into the negative range. It is of course physically impossible to have

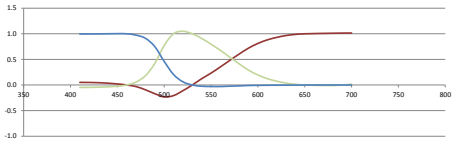
negative lights or negative photon catch at the receptor level. What this means is that some colours cannot be mixed by the arrangement of primaries as is. However, if the red light is moved onto the test half of the patch, effectively making it go on the other side of the equation $C(\lambda) + r(\lambda)R = g(\lambda)G + b(\lambda)B$, then the two patches can be matched. It is common with those experiments to use a definition of chromaticity coordinates such that by definition $\forall\lambda : r(\lambda) + g(\lambda) + b(\lambda) = 1$, so at each wavelength the sum of the three chromaticity coordinate functions is 1.

As it can be seen in Figures 4.3a and 4.3b, by using different sets of primary Guild (1931) and W. D. Wright (1928) obtained different matching functions. It is however possible to transform chromaticity coordinates from one set of primaries into another system with different primaries. The main rule is that any of the primaries cannot be obtained from a mixture of others[3]. Guild (1931) combined his own data with data from W. D. Wright (1928) by transforming his experimental results (see Figures 4.3c, 4.3d), defined in *working primaries of the instrument on which the measurements were made* into the N.P.L (National Physical Laboratory) reference system. This system is based on primaries $R = 700\,nm$, $G = 546.1\,nm$ and $B = 435.8\,nm$. When all primaries have equal chromaticity values, the colour matches with the N.P.L white point. The N.P.L is an old (read obsolete) white point illumination standard broadly equivalent to black body radiator of 4800 K. The energy distribution tabulated in Guild (1931) is normalized so that $E = 100$ at $\lambda = 560$, the same convention used for the D65 (Figure 4.1) and other standard illuminants.

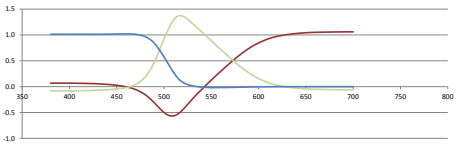
The data were averaged to describe the results of a standard observer (see Figure 4.3d). These functions were modified and used as a standard by the Commission Internationale d'Eclairage (CIE) in 1931. The new standard used the same primaries but a different white point definition. A comparison can be made between the different colour matching results in Figure 4.3 from the original data of Guild (1931) and W. D. Wright (1928) to the CIE 1931 colour matching functions. The values obtained from the colour matching experiments $(r(\lambda), g(\lambda), b(\lambda))$ are referred to as chromaticity coordinates. The chromaticity coordinates are often represented on a chromaticity diagram (Figure 4.3f) showing only two axis (usually red and green). The outline of the figure on a chromaticity diagram is composed of monochromatic colours. Coordinates outside the outline are imaginary (i.e. impossible to perceive).

As it stands, it is not sufficient to have these chromaticity coordinates to represent any colour. As Wyszecki and Stiles (1982) point out, this would not need any modification if the amount of primaries (i.e. necessary to make a match) would have been expressed in radiant power. However, the original data from W. D. Wright (1928) and Guild (1931) contained the relative brightness of the primaries. The final standard CIE 1931 of colour-matching functions, now expressed as $\overline{r}(\lambda), \overline{g}(\lambda), \overline{b}(\lambda)$, can be seen in Figure 4.4a. These functions define the standard colour matching behaviour of an observer. The

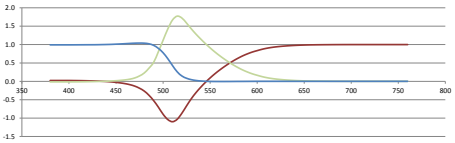
[3]This principle extends to any non-monochromatic primaries based display such as CRT guns or LED based screens.



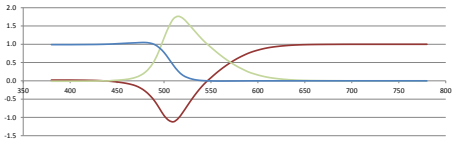
(a) W. D. Wright (1928) averaged colour-matching functions $r(\lambda), g(\lambda), b(\lambda)$ for monochromatic test $\lambda[410:700]$.



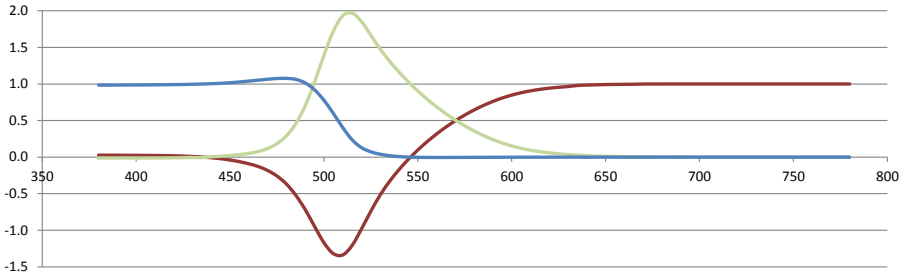
(b) Guild (1931) averaged colour-matching functions $r(\lambda), g(\lambda), b(\lambda)$ for monochromatic test $\lambda[380:700]$.



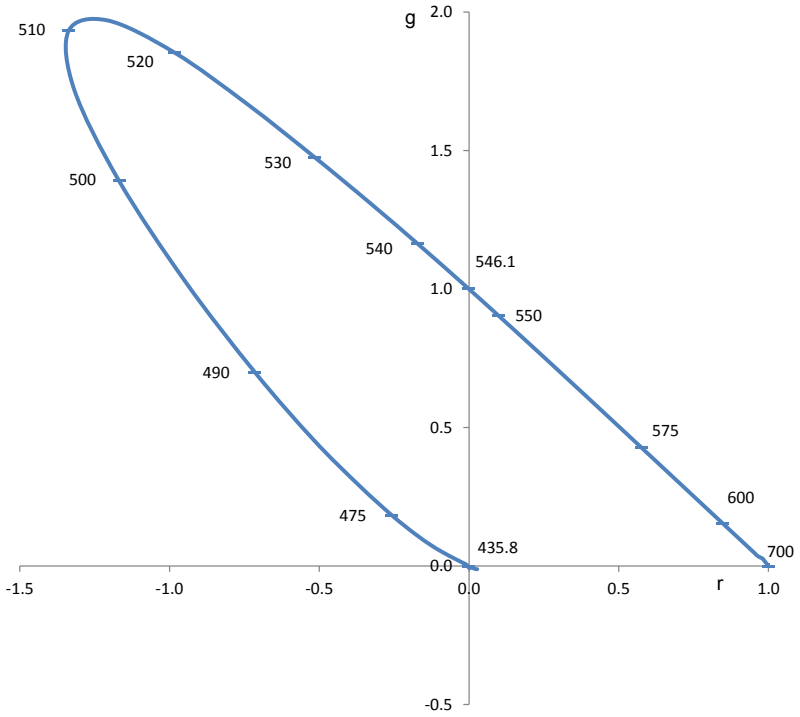
(c) Guild (1931) averaged colour-matching functions converted into N.P.L standard (see text).



(d) Averaged colour-matching functions between Guild (1931) and W. D. Wright (1928) (N.P.L standard). The experimental results are similar between the two.

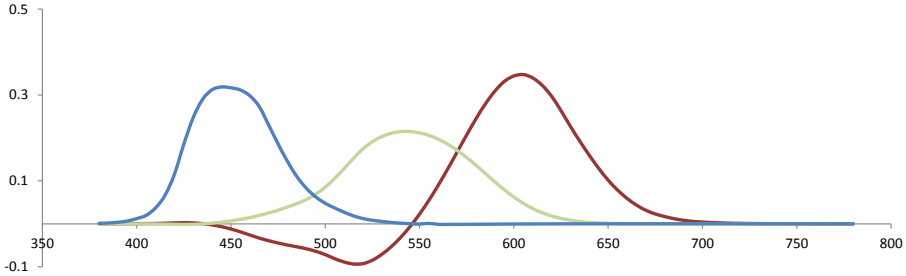


(e) Colour-matching results of standard observer as defined by the CIE 1931 standard.

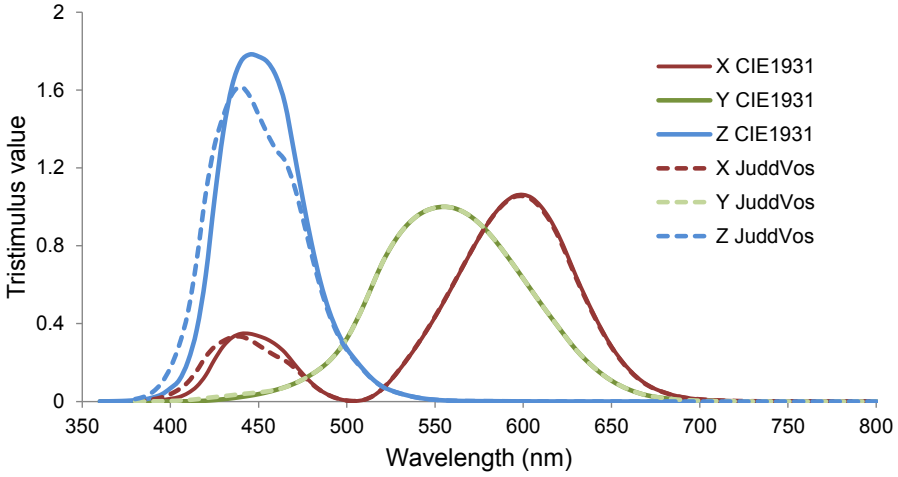


(f) Chromaticity diagram derived from chromaticity coordinates.

Figure 4.3: Chromaticity functions derived from W. D. Wright (1928) and Guild (1931) that lead to the CIE 1931 standard observer CIE (1932). The x-axis for figures (a-e) is the monochromatic wavelength (λ) the y-axis represents the chromaticity coordinates of $r(\lambda)$ (red), $g(\lambda)$ (green) and $b(\lambda)$ (blue) primaries.



(a) Colour-matching functions $\overline{r}(\lambda), \overline{g}(\lambda), \overline{b}(\lambda)$



(b) Colour-matching functions $\overline{x}(\lambda), \overline{y}(\lambda), \overline{z}(\lambda)$, original CIE 1931 standard and Judd-Vos modified.

Figure 4.4: Colour-matching functions standards

values obtained from colour-matching functions $\overline{r}(\lambda), \overline{g}(\lambda), \overline{b}(\lambda)$, are tristimulus values.

The CIE 1931 standard observer is a modified version of the RGB. The modified set is called XYZ. The choice of primaries was made to have only positive chromaticity values and so that the colour matching functions are all positive. These primaries are imaginary. The colour-matching functions $\overline{x}(\lambda), \overline{y}(\lambda), \overline{z}(\lambda)$ are are obtained from a linear combination of the original $\overline{r}(\lambda), \overline{g}(\lambda), \overline{b}(\lambda)$(see Figure 4.4). The $\overline{y}(\lambda)$ matching function was selected so that it matches the luminosity efficiency function (see Section 4.2.4, Figure 4.2). This has some obvious practical implications. We can obtain the tristimulus values XYZ of any light spectrum by integrating the radiance power over all frequencies in a similar fashion as we have seen previously for the luminosity functions (Equation 4.2).

$$
\begin{aligned}
X &= k \int_{\lambda} P_{\lambda} \overline{x}(\lambda) d\lambda \\
Y &= k \int_{\lambda} P_{\lambda} \overline{y}(\lambda) d\lambda \\
Z &= k \int_{\lambda} P_{\lambda} \overline{z}(\lambda) d\lambda
\end{aligned}
\tag{4.3}
$$

P_λ is the monochromatic radiant power at wavelength λ.

By definition $\overline{y}(\lambda)$ is $V(\lambda)$, the luminous efficiency function. Consequently if P_λ, in Equation 4.3, is defined in the same unit as $L_{e,\lambda}$, in Equation 4.2 and $k = 683$, then Y represents the luminance of the light used in cd/m^2 (see Section 4.2.4). X and Z represent the chromaticity properties of the light. The tristimulus value are commonly transformed into $x = X/(X + Y + Z)$, $y = Y/(X + Y + Z)$, $z = Z/(X + Y + Z) = 1 - x - y$, the chromaticity coordinates. A common representation for colour, as we have seen with the RGB coordinates (Figure 4.3f), is to use the xy chromaticity diagram. As z is redundant, i.e. it can be derived from x and y, a common notation is to use the xyY representation, which gives complete chromaticity and luminance information. The CIE 1931 is still widely used as an industry standard; however, the CIE has refined the official standard on consideration of the latest research in colourimetry and photometry.

Specific devices can be used to measure directly what I have described in this section and the previous one on luminance. A luminance meter will give output in cd/m^2 and is sufficient for a monochromatic experiment. In order to measure the chromaticity of an object one can use a colourimeter, a device that will measure the chromaticity coordinates and luminance. It is relatively affordable and is widely used in vision science. However, for colour experiments requiring a more precise approach, it is recommended that a spectroradiometer be used. This device measures the spectral radiance of a light, which can be transformed into any colour system (see for example Equation 4.3).

4.2.6 Cone space

Cone space functions are another type of colour-matching function (CMF, see Figure 4.4). However, these functions are meant to reflect the absorbance of retinal cones, taking into consideration lens absorption and macular pigment. The V. C. Smith and Pokorny (1975) cone fundamentals are defined as a linear transformation of the $\overline{x}, \overline{y}, \overline{z}$ Judd (1951) and Vos (1978) modified colour matching functions (see Figure 4.4b). Any set of CMFs can be linearly transformed from one to another, the details of this transformations are beyond the scope of this introduction (for an extensive review on colourimetry, see Brainard et al., 2000 and Wyszecki & Stiles, 1982). For example, Brainard, Pelli, and Robson (2002) provide a transform between cone excitation levels in V. C. Smith and Pokorny's (1975) normalized cone fundamentals and xyY CIE 1931 values. The V. C. Smith and Pokorny (1975) cone absorption function (also called cone fundamentals) are shown in Figure 4.5 along with a more recent (and accurate) set by Stockman and Sharpe (2000). The three types of cones are referred to as L, M and S for long, medium and small wavelength absorption spectra. The terms red, blue, and green cones are obsolete and are not used anymore to avoid misconceptions on the actual processing of colour.

It is of interest to note that the luminosity function (Section 4.2.4, Figure 4.2) lies between the L and M cone sensitivity functions which would suggest that luminance

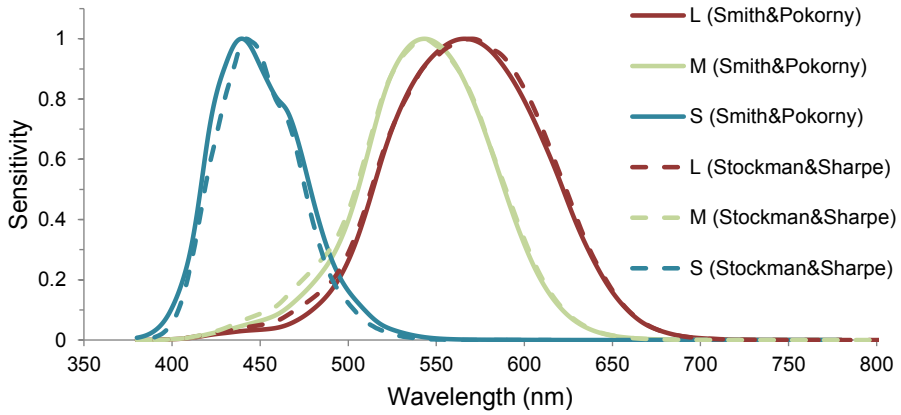


Figure 4.5: Solid line: Cone absorption functions as defined by V. C. Smith and Pokorny (1975). Dashed line: Cone absorption functions as defined by Stockman and Sharpe (2000).

perception is not computed by one cone class only but mostly by a combination of L and M cones.

4.2.7 Contrast

It is usually thought that the visual system encodes information in terms of contrast (Shapley & Enroth-Cugell, 1984). The contrast is a measure of variation of intensity. The contrast has several mathematical definitions with two definitions predominantly used in vision science: Michelson contrast, Michelson (1927) and the Weber contrast. The Michelson contrast is calculated using the highest and lowest intensity within an image (visual array or visual input), respectively denoted I_{max} and I_{min}. Usually measured as luminance ($cd.m^{-2}$, see Section 4.2.4), it is defined by:

$$M_{contrast} = \frac{I_{max} - I_{min}}{I_{max} + I_{min}} \tag{4.4}$$

Contrast can also be described with respect to a background value (Weber contrast). The Weber contrast of a test patch of intensity I_{test} with respect to a background of intensity I_b is defined by:

$$W_{contrast} = \frac{I_{test} - I_b}{I_b} \tag{4.5}$$

Following this definition, we can have negative and positive contrast values if the patch is of higher or lower intensity than the background, as opposed to the Michelson contrast, which ranges from 0 to 1 (however, it is common practice to report the Michelson contrast as a percentage). Michelson contrast tends to be used with recurrent patterns such as sine-waves (sinusoids) and square-waves. For these types of patterns, the difference between darker and lighter peaks (what would be described as negative and positive values using Weber's contrast metric) is equal. Weber's contrast is usually used only to describe single

patches on screen.

These contrast metrics have limitations. For example, more complex stimuli than single patches and recurrent patterns might require more advanced metrics (Peli, 1990); one could use root-mean-square contrast (RMS contrast) instead (see following Section). Regarding colour perception, using the appropriate contrast metric is also a complex subject. For experiments with monochromatic stimuli, one must only be concerned with reporting the background colour (usually in xyY CIE values) and one value of contrast is enough to capture the information to reproduce the experiment on another setup. Chromatic experiments might require a more elaborate description of contrast. To remedy this issue, traditionally papers describing chromatic experiments reported CRT gun (red, green, and blue) luminance modulations independently (e.g. see classic paper by Mullen, 1985, on colour channels). In recent years this has been replaced by dedicated colour contrast metrics such as cone contrast (e.g. see Mullen, Thompson, & Hess, 2010).

Cone contrast

Cone contrast is a measure of contrast that is not based on stimulus intensity (i.e. light) per se, but based on the response intensity of a cone class. Cone contrast as defined by Cole and Hine (1992) and Cole, Hine, and McIlhagga (1993) is a variation on Weber's contrast (defined earlier in this section). Each of the three dimensions of this colour space (L_c, M_c, S_c) is computed by dividing the cone intensity by the background intensity. Cone contrast T_C is defined for each type of cone ($T = L$, M *or* S) by Equations 4.6:

$$\begin{aligned} L_C &= \frac{\Delta\epsilon_L}{\epsilon_{bL}}, \\ M_C &= \frac{\Delta\epsilon_M}{\epsilon_{bM}}, \\ S_C &= \frac{\Delta\epsilon_S}{\epsilon_{bS}} \end{aligned} \tag{4.6}$$

where $\Delta\epsilon_L, \Delta\epsilon_M, \Delta\epsilon_S$ is the variation of the stimulus cone excitation from the background cone excitation, $\Delta\epsilon_c = \epsilon_{tc} - \epsilon_{bc}$, where ϵ_{tc} is the absolute cone excitation value of cone c at the test patch and ϵ_{bc} is the background excitation value of that cone type. Note that, as is the case for the Weber contrast, cone contrast can take positive and negative values.

4.3 Stimuli generation

This section outlines the methodology used to produce the stimuli used in the experimental chapters of this thesis.

4.3.1 Sinusoid and square-wave stimuli

This section describes the method used to create chromatic (red-green) or luminance recurrent stimuli, sine-waves and square-waves of a specific spatial frequency at a specific orientation. The stimuli are described in term of cone contrast, defined in the section above. In this section, I will also describe how this modulation relates to Michelson contrast.

Recurrent stimuli defined in cone contrast space

In the experiments presented in this thesis, I used sinusoidal (and square-wave) stimuli defined in cone contrast. The variation of cone contrast (T_C, as defined from Equation 4.6) of cone type T ($T = L,\ M\ or\ S$) for a sinusoidal grating is given by the following set of equations:

$$L_{Csin}(x,y) = (\pm A_L)\sin([(ycos\theta) - (x\sin\theta)] * 2\pi f) \tag{4.7}$$

$$M_{Csin}(x,y) = (\pm A_M)\sin([(ycos\theta) - (x\sin\theta)] * 2\pi f) \tag{4.8}$$

$$S_{Csin}(x,y) = (\pm A_S)\sin([(ycos\theta) - (x\sin\theta)] * 2\pi f) \tag{4.9}$$

In Equations 4.7, 4.8 & 4.9, x and y represent the relative horizontal and vertical distance from the stimulus centre in visual angle. The parameter θ corresponds to the orientation of the sinusoid, $\theta = 0°$ produces a vertical sinusoid, $\theta = -45°$ a left oblique sinusoid and $\theta = +45°$ a right oblique one. [4] The parameter f represents the spatial frequency (in cycles per degree). The parameter A corresponds to the maximum modulation of the cone contrast. If $A = 0$ (as it is for the S cone in the red-green chromatic component), then there is no modulation of contrast for this cone; the cone excitation will stay at the background value (Equation 4.6 becomes $\Delta\epsilon_C = C_C * \epsilon_{bC}$, and the result of Equation 4.7 replaces C_C).

The sinusoids can be transformed into a square-wave using the following set of equations:

$$L_{Csq} = \begin{cases} max(L_{Csin}) & \text{if } L_{Csin} > 0 \\ min(L_{Csin}) & \text{if } L_{Csin} < 0 \end{cases} \tag{4.10}$$

$$M_{Csq} = \begin{cases} max(M_{Csin}) & \text{if } M_{Csin} > 0 \\ min(M_{Csin}) & \text{if } M_{Csin} < 0 \end{cases} \tag{4.11}$$

$$S_{Csq} = \begin{cases} max(S_{Csin}) & \text{if } S_{Csin} > 0 \\ min(S_{Csin}) & \text{if } S_{Csin} < 0 \end{cases} \tag{4.12}$$

So far I have detailed how to create sinusoids and square-wave in cone contrast, using

[4] For clarity the orientation values are expressed in degrees instead of radians.

A_T as a measure of cone contrast modulation for each types of cones. The amplitude A_T is defined by:

$$A_T = a + b \tag{4.13}$$

(see Equations 4.7, 4.8 & 4.9), the value of b is used as modulation to accommodate for differences in equiluminance point (see Chapter 8 on Equiluminance and physiology Chapter 2) and only used in the context for chromatic stimuli. The values of b are determined experimentally for each participants and each value of a; this is only the case for chromatic stimuli. The chapter on Equiluminance (Chapter 8) details how equiluminance can be measured and presents the results. More details about colour contrasts and isoluminance can be found there. In this document, when I refer to a stimulus contrast I refer to the amplitude of cone contrast modulation before modulation, i.e. parameter a. In the case of luminance modulation, this parameter is equivalent to A_L, A_M and A_S and is equal to the Michelson contrast. The next section presents formally how luminance and chromatic stimuli are created and the following section will present the formal proof that the modulation parameter A is equivalent to the Michelson contrast.

Luminance and red-green stimuli

Depending on how the relative cone contrasts are modulated, it is possible to create luminance, chromatic and even combinations of both modulations. However, I detail here how to create the two separately, I will detail how to create compound stimuli, i.e. two stimuli of different orientation or type (chromatic/luminance), in a latter section of this chapter.

In order to define a purely luminance modulated stimulus; I modulate each cone class in phase, using the same contrast intensity. L, M, and S contrasts were always modulated with equal amplitude, i.e. $A_L = A_M = A_S$.

In order to create a red-green (chromatic) modulation, the cone contrasts were modulated by having L and M contrast modulated 180° out of phase, initially ($||A_L|| = ||A_M||$). In my implementation, this is done by giving A_L and A_M opposite signs. Concretely in the code $A_M = -A_M$ and $A_S = 0$. Equivalently it is possible to add a phase parameter to the right side of equations 4.7, 4.8 & 4.9.

In producing chromatic stimuli, it is necessary to make them isoluminant to the individual. Generally, the contribution of L and M inputs to the luminance signal are not equal (Gunther & Dobkins, 2002). Consequently, when L and M cone contrasts are out of phase the stimulus might not be equiluminant. In order to correct for this and create equiluminant stimuli, we adjusted the amplitude values. This is performed by adjusting both A_L and A_M. These are adjusted by the same amount keeping the sum of the two values constant (more details can be found in the Equiluminance chapter 8).

This adjustment corresponds to the value b in Equation 4.13, this parameter is different for each participant and is dependent on the amount of contrast modulation amplitude

(a in Equation 4.7). A control experiment to set equiluminance for each participant was therefore required for both sinusoids and square-wave stimuli.

4.3.2 Summary of stimuli production

Luminance or chromatic sinusoids were created using Equations 4.7, 4.8 & 4.9. If the stimulus was a square-wave the values were transformed using Equations 4.10, 4.11 & 4.12. This ensured that the spatial features of the stimuli are preserved between the two patterns, as this transformation preserves orientation, size, spatial frequency and contrast[5].

After obtaining the spatial surface of cone contrast to display on screen, I transform cone contrast values into absolute cone excitation values (ϵ_L, ϵ_M & ϵ_S). The cone excitation value (ϵ_v) is calculated by using the definition of cone contrast transformed (here shown for cone type $T = L, MorS$).

$$C_T = \frac{\Delta\epsilon_T}{\epsilon_{bT}} \tag{4.14}$$

$$C_T = \frac{\epsilon_{vT} - \epsilon_{bT}}{\epsilon_{bT}} \tag{4.15}$$

$$\epsilon_{vT} = \epsilon_{bT} + (C_T * \epsilon_{bT})$$

$$\epsilon_{vT} = \epsilon_{bT}(1 + C_T) \tag{4.16}$$

Once the three surfaces of excitation values (i.e. for L, M & S) are known, the values are converted into gun values using a matrix transform, as described by Brainard et al. (2002). This matrix transforms any triplet of excitation values into gun values bounded between 0 and 1. If the values obtained are outside of [0,1] then the contrast values that were put in are outside of the gamut of the screen.

The construction of the matrix transform from LMS cone excitation (V. C. Smith & Pokorny, 1975 cone excitations $\epsilon_L, \epsilon_M, \epsilon_S$) onto gun values specific to the screen used is described in Brainard et al. (2002). The methodology involves recording the emission spectra of the individual guns (red, green & blue) at the highest intensity. The gun spectra were obtained by using a PR-650 (Photo Research) spectroradiometer.

Before being displayed, the latest step is the gamma correction of screen non-linearity, also described in Brainard et al. (2002). The gamma correction procedure is detailed in a following section.

In this section, I have detailed how to create luminance and chromatic recurrent stimuli. Two types of stimuli were detailed, sinusoids and square-waves. The procedure from contrast modulations to gun values was reported. The next section will detail how to display the stimuli, specifically relating to the technique of gamma correction and frame

[5]The contrast of square-wave versus sine-wave is a more complex issue and is explored in the first experimental Chapter 5, what is meant here is that highest and lowest intensity values are preserved and as most definition of contrast only uses those two values (see Weber and Michelson contrast in this Chapter), then the value of this contrast metric is preserved, as we will see in the experimental Chapter the actual perceived contrast might be different.

interleaving, as well as detailing how to display compound stimuli (i.e. two stimuli generated with the method described above but with different parameters).

4.3.3 Display technique

Gamma correction

It is known that the luminous intensity of CRT guns does not follow a linear relationship with the voltage used to control them. This voltage is abstracted from 0 to 1 for the user.

All three guns tend to follow a gamma function as a function of voltage. As the chromatic space is calibrated at the highest values and assumes linearity this is problematic. However, there is a common solution, which is to use the inverse, symmetrical function and change the value that is asked to be displayed to match that of the one obtained by the inverse function.

The three gamma functions are obtained empirically by measuring luminous intensity at constant intervals from 0 to highest voltage. This was done by an automatic routine in the CRS Visage, using a ColorCal colorimeter. However, the 64-bits mode display on the ViSaGe does not allow automatic gamma correction consequently, I used an in-house Matlab function to fit the data collected and to create an inverse function. The gamma correction function was used as the last step to display stimuli on screen. The gamma correction routine was run regularly in the lab to update the functions.

Frame interleaving

Most stimuli described in this thesis are actually compound stimuli of luminance and chromatic components, as I am interested in the perceptual interactions between the two. So far, I have described how to generate individual frames with either luminance or chromatic variations applied to them. However, I need to be able to display overlaying stimuli of different contrast, different orientation and different type (luminance/ chromatic).

In order to have both variations it would be possible to create two stimuli and sum up the two or update the stimuli definition. However, this risks encountering screen interaction non-linearities.

The common solution for this is to use frame-interleaving (e.g. see Ross & Speed, 1991). Using a stimulus consisting of two frames, one carrying the luminance stimuli and the other the chromatic (or for a detection task with mask, one frame with the target and one frame with the mask with independently adjustable contrasts). If the frame rate is high enough (and this was the case for all setups used in this thesis) then the fusion becomes seamless.

Frame interlacing of a chromatic component is common in colour vision psychophysics (e.g. Victor, Purpura, & Conte, 1998; Michna, Yoshizawa, & Mullen, 2007; Kingdom et al., 2005). This technique has the advantage that it limits any non-linear interactions on the screen (Victor et al., 1998). The drawback of this method is that it effectively decreases the actual contrast by a factor of two. All contrast values reported in this thesis

take into account this fact and are actual contrast and no single frame contrast. The high frame rate of the monitors used in my experiments (110 Hz and 100Hz) allows seamless fusion for the observers.

In the next section, I will explore a bit further the contrast properties of the stimulus generated by the method described in the above sections. I will detail further how the metric used A_T relates to other contrast metrics.

4.3.4 Stimulus contrast

This section focuses on the relation between the contrast amplitude A, which was defined previously as a modulation of cone contrast across space, and other popular contrast metrics. In the first subsection, I will demonstrate that this metric is equivalent to the Michelson contrast. Secondly, in a following subsection I will detail how to calculate the R.M.S cone contrast from cone amplitude modulation.

Michelson contrast

For this demonstration I will use the intensity terms I_L as in the equations of contrasts (Equations 4.5) but as this term directly transfers into cone intensity as described by Cole et al. (1993) (see Equation 4.6) for cone contrast, this demonstration therefore generalizes to cone contrast.

In Equations 4.7, 4.8 & 4.9, detailing the sinusoidal modulation of cone contrast, sin and cos functions produce periodic waveforms oscillating from -1 to $+1$. In the left-hand side of Equations 4.7, 4.8 & 4.9 we multiply this output by $\pm A$ (in the luminance case $b = 0$). Hence there is an oscillatory wave going from $+A$ to $-A$ in cone contrast space across the stimulus. This gives a total amplitude peak-to-peak of $2A$.

As the Michelson contrast is concerned with highest and lowest intensity values (Equation 4.4), I defined the highest contrast peak as $C_{max} = +A$ and the lowest as $C_{min} = -A$. I get positive and negative contrast; this is not surprising using a definition of contrast based on Weber contrast (Cole et al., 1993; Equation 4.5). I then turn these values into intensity (I_{test}) using Equation 4.5, knowing the intensity of the background (I_b) and the contrast of the test (C_{test}).

$$\begin{aligned} C_{test} &= \frac{I_{test} - I_b}{I_b} \\ I_{test} &= I_b + (C_{test} * I_b) \\ I_{test} &= I_b(1 + C_{test}) \end{aligned} \tag{4.17}$$

As stated previously, there are two intensity values of interest, the highest (I_{max}) and lowest (I_{min}) in order to use the Michelson's contrast equation. Using the result of Equation 4.17, I can get the highest and lowest cone contrast values (recall that $C_{max} = +A$ and $C_{min} = -A$) and transform them into intensity values. Consequently I get

$I_{max} = I_b * (1 + C_{max}) = I_b * (1 + A)$ and $I_{min} = I_b * (1 + C_{max}) = I_b * (1 - A)$. Now that I have both I_{max} and I_{min}, I can calculate the Michelson contrast (using Equation 4.4).

$$\begin{aligned} M_{contrast} &= \frac{I_{max} - I_{min}}{I_{max} + I_{min}} \\ &= \frac{(I_b * (1 + A)) - (I_b * (1 - A))}{(I_b * (1 + A)) + (I_b * (1 - A))} \\ &= \frac{(I_b + I_b * A) - (I_b - I_b * A)}{(I_b + I_b * A) + (I_b - I_b * A)} \\ &= \frac{(I_b + I_b * A - I_b + I_b * A)}{(I_b + I_b * A + I_b - I_b * A)} \\ &= \frac{I_b * (1 + A - 1 + A)}{I_b * (1 + A + 1 - A)} \\ &= \frac{I_b * 2A}{2 * I_b} \\ &= A \end{aligned} \tag{4.18}$$

This is the same proof for cone contrast if we use the notation in Equation 4.6 instead of the Weber definition in Equation 4.17. Replacing luminance intensity (I) by cone excitation for three cone classes (ϵ_c).

R.M.S cone contrast

Gunther and Dobkins (2002) presented an equation to measure the Root Mean Square (R.M.S) cone contrast for chromatic stimuli. This definition is given for red-green chromatic stimuli. As for cone contrast, this definition uses cone excitation values. In this original paper, they describe values in term of red peak and green peak. This actually corresponds to maximum cone excitation and minimum cone excitation for L cone and inversely for M cone.[6] The equation is as follows:

$$\sqrt{\frac{[(\epsilon_{Lmax} - \epsilon_{Lmin})/(\epsilon_{Lmax} + \epsilon_{Lmin})]^2 + [(\epsilon_{Mmin} - \epsilon_{Mmax})/(\epsilon_{Mmin} + \epsilon_{Mmax})]^2}{2}} \tag{4.19}$$

Additional comments on contrast metrics can be found in the Equiluminance chapter 8. The equiluminance chapter also details the whole methodology to create equiluminant stimulus as well as present the equiluminant data obtained from each main experiment presented in this thesis. In this chapter, I also discuss the impact of modulation of contrast amplitude (A_L and A_M) toward contrast metrics in the case of chromatic stimuli.

[6] Consequently, the notations L_r, L_g, M_r *and* M_g in the original paper Gunther and Dobkins (2002) are replaced by $\epsilon_{L,max}, \epsilon_{L,min}, \epsilon_{M,min}$ *and* $\epsilon_{M,max}$.

4.4 Summary

I have detailed in this section how to stimulus can be described using several contrast metrics from cone to luminance. I have also presented a brief summary of how the luminance and chromatic matching functions were obtained. Next, I have shown how to construct the stimuli used in this thesis by detailing their mathematical definition and explaining the display techniques used.

Additional methods regarding the creation of isoluminant stimuli, through isoluminant experiment (such as minimum motion and minimum flicker) are presented in a dedicated chapter (Chapter 8). This chapter also contains the results of associated isoluminant experiments and discussion regarding contrast metrics and cone ratio.

Chapter 5

Perception of shape[1]

5.1 Aims

The aim of the study described in this chapter was to explore the generality of the heuristics proposed by Kingdom (2003), described in the Introduction (see Section 3.6.6, p. 65) for the interpretation of shape from shading and colour, by testing if the colour-shading effect works with different kinds of shape profiles. A straightforward manipulation is to replace sinusoids with a hard-edged pattern, such as a square wave on screen (producing a triangle-shaped profile in depth). This effectively adds some harmonics of spatial frequency (Kingdom & Simmons, 1998). In terms of the shape perceived it turns a corrugated pattern that is close to sinusoidal in depth (Kingdom et al., 2005; M. J. Wright & Ledgeway, 2004) into a hard-edge folded triangular wave. It has been shown by P. Sun and Schofield (2012) that square wave luminance patterns tend to be perceived as triangular surfaces in depth.

Another important point of this study was to simplify the experimental design and procedure that had been used before. In previous studies (Kingdom, 2003; Kingdom et al., 2005, 2006), stimuli consisted of three components: the colour sinusoidal component was always paired with a luminance component of the same orientation. Experiments explored how those luminance-colour pairs affected the depth perceived in a third luminance-only or colour only sinusoidal component (which could have a range of orientations). Yet the logic behind the interpretation of the effects found suggests that the effect should work with only a pair of components. Here I used only two oriented components, one red-green chromatic, and one luminance-defined. This was done to avoid possible contamination of the results from low-level masking effects, which have been described between luminance and colour (e.g. Medina & Mullen, 2009).

In this chapter, I will test the hypothesised heuristics presented by (Kingdom, 2003) regarding the extraction of reflectance versus the extraction of shape. I will test the generality of this heuristic in two ways. First, I will use a simplified design with only two

[1]Parts of this chapter have been published, the publication, (Clery, Bloj, & Harris, 2013), is attached in Appendix B.

Figure 5.1: a) Top: Illustration of the matching stimulus used for the experiment. When presented via a stereoscope, observers perceive a circular depth pedestal in the middle of the matching stimulus. Bottom: colour shading stimulus, always identical in both eyes (no disparity). b) Side view of the disparity profile, the inner disk of the top stimulus is adjustable using a dial. c) Depth profile, observers were asked to match the depth perceived from shading to the disparity-defined stimulus.

components (one luminance and one chromatic) instead of the three-component design used in the original study. If the colour-shading effect is indeed related to the extraction of reflectance versus shape, then it should also work with a simpler configuration of the stimulus. Secondly, I will introduce two types of patterns, a sinusoid pattern (as a reproduction of the original study) and a square-wave pattern. The square-wave pattern introduces edges in the final percept; these edges once again could be interpreted to be an abrupt change in shape (i.e. a fold) or an abrupt change in reflectance (or colouring) of the surface. The change of pattern and the use of a simpler design (two components only) will help to assess the generality of the heuristic hypothesized to be the reason for the colour-shading effect.

5.2 Part 1

5.2.1 Methods

Observers

A total of 12 observers were tested. All observers were naïve to the task and hypothesis tested. All of them had no colour deficiency as verified by the Ishihara colour plate test (38-plate edition, 1979). These observers were also screened with a standard TNO test for disparity perception deficiencies. All observers had normal or corrected to normal acuity.

Ethical approval was given by the University of St. Andrews ethics committee UTREC (reference PS6135) and followed university guidelines. Participants were given monetary

compensation for their time.

Apparatus

The stimuli were displayed on a CRT colour monitor Mitsubishi Diamond Pro (110 Hz, non-interlaced). The CIE-1931 chromaticity coordinates of the red, green, and blue phosphors were $x = 0.620, y = 0.340; x = 0.290, y = 0.604$ and $x = 0.149, y = 0.071$, respectively. The gun chromaticity values were obtained by measuring the full spectra of the gun with a Photo Research spectroradiometer PR-650 and then converting into CIE-1931 (2° Standard observer) values by multiplying the functions and spectra. The stimuli were generated using Cambridge Research Systems ViSaGe and Matlab. The screen gamma non-linearity was corrected using a CRS ColorCal colorimeter. Details on stimulus generation as well as the colour space used are described in the following sections.

Participants viewed the screen through a Wheatstone stereoscope. Using sets of mirrors, this stereoscope splits the available visual field on the CRT in two. Effectively the left part of the screen will be only visible to the left eye and the right part will be only visible to the right eye. A chin rest was mounted with the stereoscope to hold the observers head steady at the chosen viewing distance of 70 cm.

Stimuli and task: General

Stimuli consisted of combinations of luminance and chromatic sinusoids or square waves, positioned below a random dot stereogram depicting a circular slab at a different depth from its circular background (see Figure 5.1a). Below, I describe the stereo-defined stimuli, represented in the upper half of Figure 5.1a. Observers were required to adjust the stereoscopic depth-matching stimulus until it appeared to contain the same depth (defined by binocular disparity, Figure 5.1b) as the peak-to-trough depth depicted by the luminance sinusoid or square wave (defined by shape-from-shading), see Figure 5.1c.

Stimuli: depth-matching

A stereoscopically defined depth pedestal was used to assess the amount of perceived depth in the target luminance pattern. The stimulus (see top half of Figure 5.1) was a random dot stereogram, composed of a circular patch (4 deg diameter), made of pre-rendered circular Gaussian blobs (standard deviation of 0.8 mm on screen, or 3.4min arc). A central circular pedestal (2 deg diameter) could be adjusted in depth (see disparity profile Figure 5.1b). The blobs in the random dots stereogram were non-oriented Gaussian envelopes, darker than the background. The blobs with coordinates falling in the inner 2 degrees disk were modulated in disparity, i.e. when the observer adjusted a dial the stimulus was redrawn with the central patch having a new disparity. The background blobs (outer circle) were presented at 0 min arc disparity (see Figure 5.1b).

Previous experiments on the colour shading effect used disparity defined matching stimuli that resembled the final percept. As the aim here was to compare the depth

perceived via sinusoids and square wave patterns, I chose to use a disk-shaped pedestal. This matching stimulus makes no assumptions about the shape of the perceived object, which can vary for the square wave (see P. Sun & Schofield, 2012). Observers were asked to turn a dial to match the depth between the central patch and the background (Figure 5.1c) with the amount of depth they perceived from shape-from-shading for the luminance grating: peak to trough from the lowest to the highest depth.

Stimuli: luminance and chromatic

The test stimuli were composed of one chromatic and one luminance component of the same profile type, but with different orientation and contrast. The luminance component carried the shape of the object according to shape-from-shading. In this experiment, the type of pattern used to generate the chromatic and achromatic components could be either sinusoidal or square wave modulations. A square wave pattern on-screen, corresponds roughly to a triangular profile in depth. A sinusoidal pattern corresponds to a corrugated material with an approximate sinusoidal shape. All stimuli used are shown in Figure 5.2 for both sinusoids and square waves. Below I describe the methodology and colour space used to create these patterns and later the experimental conditions tested.

Both chromatic and luminance components had a spatial frequency (fundamental frequency for the square wave) of 0.75 cycles per degree. An equation describing how to generate a sinusoid can be found in Methods Chapter Section 4.3.1, as well as the modification used to obtain a square wave pattern. To avoid stimulus display artefacts, the two components (i.e. chromatic and luminance) were displayed on different video frames and the frames were temporally interleaved, see Chapter 4 General Methods.

The luminance component, which gives the impression of shape, was always oriented at $-45°$ (left oblique). The chromatic component could be either oriented at $-45°$ (left oblique, aligned with the luminance component) or $+45°$ (right oblique, orthogonal to the luminance component). These two conditions are sometimes referred to in the literature as iso-oriented and cross-oriented, respectively.

The chromatic component was obtained by setting modulations of inputs to L and M cones to be out of phase (L-M). Thus when L-cone activation is at its highest, M-cone activation is at its lowest, and vice versa. The chromatic component does not change the overall shape of the object; however, it is expected to have a modulatory effect on the perceived depth of the corrugations.

Colour space definition can be found in Chapters 4 and additional details in Chapter 8.

Experimental design and procedure

Perceived depth was measured in the shape-from-shading delivered by the luminance component, using method of adjustment. Observers used a stereo defined patch to adjust the perceived depth. Several different combinations of luminance and chromatic components

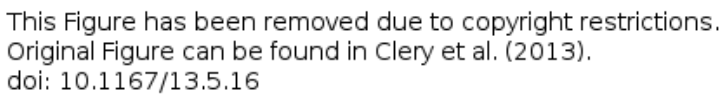


Figure 5.2: Illustration of the stimuli used. All stimuli are shown for sinusoids and square-wave. Each condition is shown on a different row. For each condition, one component is fixed and one component is systematically changed. a) Condition 1, chromatic component fixed and variable luminance contrast in orthogonal orientation. b) Condition 2, luminance component fixed and variable contrast in orthogonal orientation. c) Condition 3, luminance component fixed and variable contrast in aligned orientation. During experimental trials, conditions were all randomized and sinusoids and square-wave were tested in different sessions.

were used, which I will describe below as three conditions. Trials from each condition were randomly interleaved and the task was always the same. The experiment was split into two sessions, one for sinusoids and one for square wave patterns (the order of which was randomized between observers). Before each session, observers performed a minimum motion experiment with identical stimuli to those used during the main experimental sessions. The results of the minimum-motion as well as the design of the experiment are presented in the Isoluminance Chapter 8.

Observers performed 10 trials per contrast value for each condition and for each pattern profile. Figure 5.2 show an illustration of all stimuli for both the sinusoid and the square wave condition.

Condition 1

In order to measure the effect of increasing luminance contrast on perceived depth, the colour component was fixed with a right-oblique orientation (+45°) at a constant contrast amplitude of 0.012. The luminance component was left-oblique oriented (+45°), and tested using six values of contrast (0, 0.01, 0.02, 0.04, 0.08 0.16; see Figure 5.2a). In this case, the luminance contrast measure is equivalent to Michelson contrast (Methods Chapter Section 4.2.7) and so these values can be expressed as percentages.

Condition 2

The luminance left-oblique component was fixed at contrast amplitude 0.04 and six values of colour contrast were tested: 0, 0.004, 0.08, 0.012, 0.016, 0.02 (see Figure 5.2b). It was expected, that the perceived depth would increase as the colour contrast increased (see, Kingdom, 2003; Kingdom et al., 2005, 2006). Note that the fourth stimulus of condition 2 is exactly the same as the third stimulus of condition 1 (0.04 luminance component and orthogonal chromatic component at 0.012 contrast amplitude) as can be seen in Figure 5.2:a-b.

Condition 3

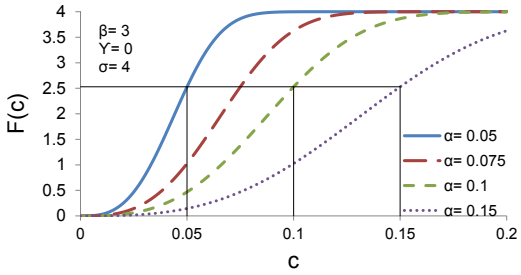
The colour component was left-oblique oriented and hence aligned with the luminance grating. The values of contrast used were the same as in condition 2 (Figure 5.2c). Note that condition 2 and 3 are indistinguishable at 0 colour contrast, this gives us the perceived depth for luminance only at a fixed contrast of 0.04. From previous work (Kingdom, 2003; Kingdom et al., 2005, 2006), the predication was that aligned in-phase chromatic contrast would suppress perceived depth.

Data Analysis

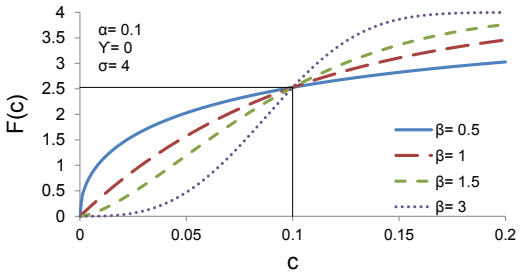
I analysed sinusoid and square-wave data conditions separately. Matched-depths were recorded as a function of contrast for each of the three conditions separately, for each observer. A function was then fitted to the matched-depth versus contrast data, separately for sinusoids and square waves, this was contingent on showing a significant effect of contrast on depth, as I will describe latter on.

It is common for luminance contrast mechanisms to show sigmoidal/saturation behaviour (Albrecht & Hamilton, 1982; Dean, 1981; Schofield, Rock, Sun, Jiang, & Georgeson, 2010) and this has been hypothesized to linked to normalizing mechanisms (see Chapter 2 contrast saturation, see also Carandini & Heeger, 2012). Hence, in condition 1 perceived depth data was exclusively fitted with a modified cumulative Weibull function:

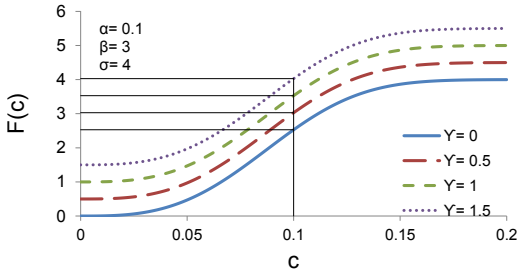
$$F(c) = \gamma + \sigma(1 - e^{-(c/\alpha)^{\beta}}) \tag{5.1}$$



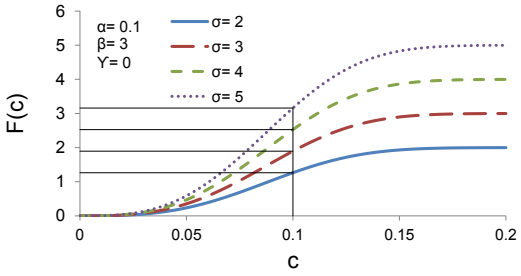
(a) α parameter, determines the threshold, defined as the 63% of increase from baseline, γ, to ceiling asymptote $\gamma + \sigma$. (Threshold is 2.53).



(b) β parameter, determines the shape or slope, higher values make the shape more sigmoidal. Note that α has a compressing effect on the slope shape (see a).



(c) γ parameter, determines the baseline asymptote.



(d) σ parameter, determines the span between ceiling and floor asymptote, in conjunction with γ it therefore determines the ceiling asymptote.

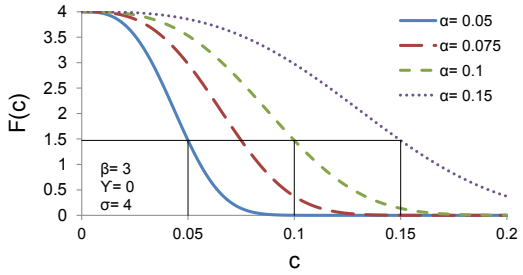
Figure 5.3: Description of cumulative Weibull function parameters

I describe here how the Weibull function behaves, Figure 5.3 illustrates how the parameters of Equation 5.1 can vary. Equation 5.1 has four free parameters γ (Figure 5.3c), the offset parameter, the scale parameter σ (Figure 5.3d), used to scale the functions output on the y-axis, α the threshold parameter (Figure 5.3a), and β the shape parameter (Figure 5.3b). High values of β give step-like functions and lower values will deliver a more sigmoidal function. We define threshold as the value of α that delivers a value for the function of 63% of maximum.

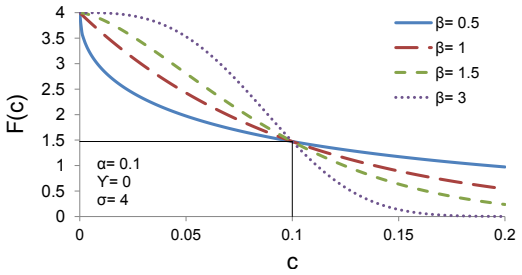
Decreasing effects are also expected in some conditions; consequently, a monotonic decreasing function is required. I used a modified version of the cumulative Weibull to do so:

$$F(c) = \gamma + \sigma(e^{-(c/\alpha)^{\beta}}) \tag{5.2}$$

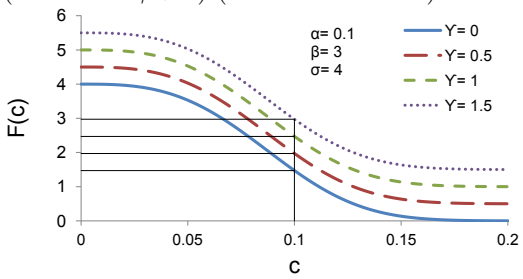
I illustrate the behaviour of the inverse cumulative Weibull in Figure 5.4. Equation 5.2 has four free parameters similar to Equation 5.1. In Figure 5.4 I used the same values of parameters to illustrate the differences and compare with Figure 5.3. γ (Figure 5.3c), the offset parameter, or floor asymptote, the behaviour is similar to the cumulative function (compare Figures 5.3c & 5.4c). σ (Figure 5.3d), the scale parameter has a similar effect, scaling the span between floor and ceiling asymptote. α the threshold parameter (Figure 5.4a), describe the point on the x-axis where the function is at 63% change from one



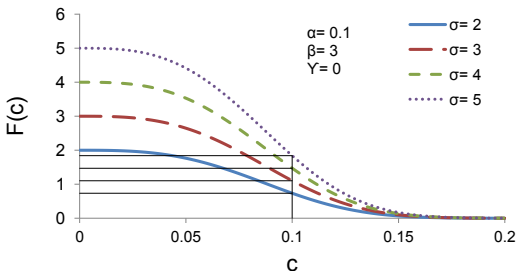
(a) α parameter, determines the threshold, defined as the 63% of decrease from ceiling (or 37% of $\gamma + \sigma$) (Threshold is 1.47).



(b) β parameter, determines the shape or slope, higher values make the shape more sigmoidal.



(c) γ parameter, determines the baseline asymptote.



(d) σ parameter, determines the span between ceiling and floor asymptote, in conjunction with γ it therefore determines the ceiling asymptote.

Figure 5.4: Description of inverse cumulative Weibull function parameters

asymptote to the other. Alternatively in the decreasing case, an equivalent definition is, α corresponds to the x-axis value when the function is at 37% of $\gamma + \sigma$ (compare 5.3a & 5.4a). β the shape parameter has the similar effect to that previously defined.

From the previous literature (Kingdom, 2003; Kingdom et al., 2005, 2006) I expected results to be monotonic either decreasing or increasing. I used an automatized fitting procedure for every participant and every condition. For each observer, pattern and condition, I entered the perceived depth data into a one-way Kruskal-Wallis test to check for a main effect of contrast (luminance or chromatic contrast, depending on condition). If there was no main effect, I did not proceed with fitting, as it would essentially be fitting noise. If the effects were significant I then performed a linear fit to determine the direction of the effect, that I describe as either positive or negative[2]. Based on the fit direction I then either fit the data with a cumulative Weibull or its decreasing counterpart, see Figures 5.3 & 5.4 and Equations 5.1 & 5.2 above.

As the luminance contrast is the main driver of depth perception (see Shape from shading section 3.6.6), I expect to have no depth perceive at 0 and then increased depth perception. The γ parameter was therefore set to 0.

For condition 2 and 3, I expected to have respectively enhancing and suppressing effects of colour contrast, as the colour will be orthogonal or aligned. The results of statistical

[2]Alternatively, I could have started by fitting a line then check if the slope parameter was significantly different from 0, I expect this method to give similar results.

analysis and fits are presented in the following section.

I report the raw results for each condition below, followed by a description and discussion of the model fits.

5.2.2 Results

Condition 1: fixed colour contrast and variable luminance component

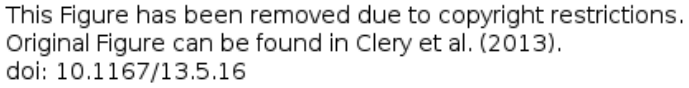


Figure 5.5: Results from condition 1: Mean perceived depth as a function of luminance contrast for each observer (black circles: sinusoids, grey squares: square-wave, error bars: standard error of the mean, SEM). The individual curve fits are presented with solid lines.

In condition 1, the independent variable was the contrast of the left oblique luminance component. The colour contrast was fixed at 0.012. The One-way Kruskal-Wallis test results are presented in Table 5.1 and the mean (± SEM) plus curve fits are presented in Figure 5.5. Mean matched depth as a function of luminance contrast is plotted in Figure 5.5 for each observer (sinusoids: black circle, square-wave: grey square). For all observers, the increase in luminance produced a significant increase in perceived depth (one-way Kruskal-Wallis test tests, $\chi^2(5,54)$, $p < 0.001$, see Table 5.1) for both sinusoids and square-wave, except for participant p5 who was unable to see the 3D shape of square

Table 5.1: Analysis results of condition 1. One-way Kruskal-Wallis test, for each participant and each pattern (sinusoid and square-wave). All effects (.i.e. slopes) where positive, when fitted with a linear regression. n.a.: not applicable; corresponds to Participant 5, who couldn't perform the task with square-wave, see text.

	Sinusoid		Square-wave	
Participant	$\chi^2(5,54)$	p	$\chi^2(5,54)$	p
P1	46.05	<.001	40.24	p <.001
P2	52.89	<.001	38.50	p <.001
P3	52.57	<.001	43.86	p <.001
P4	28.70	<.001	39.52	p <.001
P5	48.68	<.001	n.a.	n.a.
P6	47.84	<.001	45.86	p <.001
P7	46.00	<.001	54.70	p <.001
P8	50.45	<.001	48.47	p <.001
P9	48.05	<.001	40.38	p <.001
P10	46.28	<.001	46.22	p <.001
P11	42.96	<.001	41.59	p <.001
P12	49.22	<.001	46.99	p <.001

wave pattern (consequently in result tables for participant p5 in square-wave condition data is missing, indicated by *n.a.*). For all participants, similar saturating functions were obtained, this allows us to use a cumulative Weibull function (Equation 5.1, described in Section 5.2.1) to perform the fit to the data, plotted with solid lines on Figure 5.5. This is intended has data description and not a computational model of the mechanism involved.

The results are consistent with what was expect from classical shape-from-shading, where higher amplitude shading corresponds to more perceived depth. Participant 9 is the only individual with a different pattern of responses from the rest of the group. For this participant, a clear step-like function can be seen. Overall, the observers tested set maximum depths in the range 1.22 to 4.41 arcmin ($mean = 3.38, s = 0.9$).

When comparing sinusoid and square-wave data, the response to the square wave tended to reach the highest perceived depth faster than for sinusoids. These results demonstrate that all our participants were performing the task correctly, that is to say responding to the depth shown from the luminance-defined component. As all the conditions were interleaved, it can be assumed that participants were not switching tasks between conditions.

Furthermore, as sinusoids and square-waves were tested in different sessions, it is evident from the results in Figure 5.5 that, using our stimuli, shape from shading is a robust phenomenon, with similar maximum depth perceived and similar shapes of fitted function for both sinusoids and square waves. The only exception is participant p4, who shows a scaling in terms of maximum depth perceived. Seemingly, for this participant the square-wave patterns elicit more depth than the sinusoid. When looking at the fitted parameters for that participant (Table 5.2), one can see that the σ value is higher, reflecting the over-

Table 5.2: Fit results for condition 1, all conditions were fitted with cumulative Weibull and $\gamma = 0$ see text and Equation 5.1 for details.

	Sinusoid			Square-wave		
Part. ID	α	β	σ	α	β	σ
P1	0.033	0.925	2.844	0.020	0.683	2.620
P2	0.055	1.153	3.907	0.056	0.874	3.771
P3	0.040	2.157	2.254	0.029	1.096	2.127
P4	0.016	1.170	1.224	0.013	1.008	3.007
P5	0.037	1.364	4.068	na	na	na
P6	0.018	4.769	3.114	0.016	1.189	4.177
P7	0.099	1.147	2.934	0.056	1.222	2.691
P8	0.023	0.636	4.411	0.015	0.889	4.311
P9	0.040	28.908	4.328	0.020	33.457	4.253
P10	0.019	0.963	4.165	0.014	0.717	3.991
P11	0.042	1.265	2.771	0.015	3.470	2.548
P12	0.031	1.281	4.041	0.017	0.732	4.272

all higher perceived depth. However the threshold values (α) and the shape value (β) are within the same range as those found for other participants.

The results between our participants and between sinusoids and square-wave tended to be homogeneous (see the fits parameters in Table 5.2) Participant p9, showed a step-like result as translated into high β values. Furthermore regarding the fits parameters, the values of σ gives a good estimation of the maximum depth perceived with luminance contrast. Sinusoid and square-wave have similar saturating values but there are individual differences in maximum depth perceived, as noted earlier.

Condition 2: fixed luminance contrast, variable colour contrast

In condition 2, the independent variable was the contrast of the right oblique colour component, presented orthogonally to the shape-carrying luminance component. From previous research on the colour shading effect I expected to find an increase of perceived depth with increasing colour contrast (Kingdom, 2003; Kingdom et al., 2005, 2006). The One-way Kruskal-Wallis test results are presented in Table 5.3 and the mean ($\pm$ SEM) plus fits are presented in Figure 5.6. Mean matched depth as a function of luminance contrast is plotted in Figure 5.6 for each observer (sinusoid: black circle, square-wave: grey square).

They were two surprising results. First, for several participants, the chromatic component had no effect on perceived depth in the luminance component (Table 5.3 & Figure 5.6). Kruskal-Wallis test tests showed that some participants did not show any significant effect of colour contrast on perceived depth (see Table 5.3). When effects were significant, curves were fitted and plotted in Figure 5.6.

Eight participants (out of twelve) had a significant colour contrast effect ($p < 0.05$) for

Table 5.3: Analysis results of condition 2. One-way Kruskal-Wallis test, for each participant and each pattern (sinusoid and square-wave). The effect (increase or decrease) of chromatic contrast on perceived depth is obtained with a linear regression. na: non applicable, see text

	Sinusoid			Square-wave		
Part. ID	$\chi^2(5, 54)$	p	Effect direction	$\chi^2(5, 54)$	p	Effect direction
P1	15.88	<.01	increase	9.10	.105	*n.s.*
P2	0.76	.979	*n.s.*	9.00	.109	*n.s.*
P3	1.80	.876	*n.s.*	0.70	.983	*n.s.*
P4	20.74	<.001	decrease	6.22	.285	*n.s.*
P5	33.45	<.001	increase	na	na	na
P6	8.24	.144	*n.s.*	13.13	<.05	decrease
P7	10.71	.058	*n.s.*	32.95	<.001	increase
P8	6.10	.297	*n.s.*	2.51	.776	*n.s.*
P9	13.77	<.05	0.031	1.93	.859	*n.s.*
P10	3.29	.655	*n.s.*	22.99	<.001	increase
P11	4.37	.498	*n.s.*	5.13	.400	*n.s.*
P12	13.09	<.05	increase	7.97	.158	*n.s.*

Table 5.4: Fit results for condition 2, all conditions with significant effect (see previous table) were fitted with a cumulative (positive) or inverse Weibull (negative), see Equation 5.1 and 5.2 for details

	Sinusoid					Square-wave				
Part. ID	Effect	α	β	γ	σ	Effect	α	β	γ	σ
P1	increase	0.014	4.668	1.103	1.064	no fit				
P2	no fit					no fit				
P3	no fit					no fit				
P4	decrease	0.007	73.082	0.959	1.064	no fit				
P5	increase	0.006	2.831	0.532	2.604	n.a.				
P6	no fit					decrease	0.019	1.584	3.560	0.775
P7	no fit					increase	0.031	1.786	0.908	2.761
P8	no fit					no fit				
P9	decrease	0.015	4.444	1.228	2.967	no fit				
P10	no fit					increase	0.024	1.989	2.720	2.339
P11	no fit					no fit				
P12	increase	0.005	0.925	2.086	1.270	no fit				

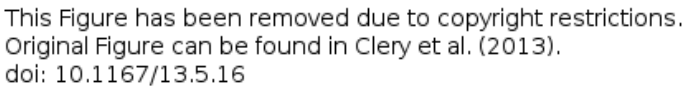


Figure 5.6: Results from condition 2: Perceived depth as a function of colour contrast for all observers. (black circle: sinusoids, grey squares: square-wave, error bars: standard error of the mean, SEM). When statistics revealed a significant effect of colour contrast (see table [5.1]), the data were fitted with a Weibull function, the individual fits are shown by solid lines.

sinusoid or square-wave or both, the other four participants showed no significant effects at all ($p > 0.05$). No participants showed a significant effect in both patterns; however, participant P1 and P12 seems to lack statistical power in the square-wave condition to become significant. Interestingly, for both of them the direction of the effect was similar (increase) in both sinusoid and square-wave, showing a consistency of effect.

The second surprising effect, and perhaps more surprising considering the literature, has to do with the direction of the effects. Of the participant with significant effects, three showed negative effects of colour contrast (P4, P6 and P9). Thus, less depth was perceived as chromatic contrast increased. This is contrary to what was reported previously in the

literature (Kingdom, 2003; Kingdom et al., 2005, 2006) and does not fit the heuristic proposed to explain the effect (Kingdom, 2003).

As stated in the Method (Section 5.2.1), the fourth contrast value in both luminance and chromatic contrast between condition 1 and 2 is the same. Essentially these conditions are tested twice. The comparison between Figure 5.5 & 5.6 shows that the fourth contrast values in both graph, for every participant and each pattern type (square-wave and sinusoids) yields similar values in term of mean and spread (SEM).

Another interesting point that can be extracted from the data in condition 2 is the baseline depth perceived at 0 colour contrast, corresponding to the first data points on the graphs (Figure 5.6). In this specific configuration, only the luminance component is visible, with a contrast of 4%. This baseline is subject to individual differences.

Condition 3: aligned fixed luminance contrast and variable colour contrast

In this condition, the same values of luminance contrast (0.04) and colour contrast were used as in condition 2, however the orientation of the colour component was set so that the chromatic and luminance components were aligned (iso-oriented and in-phase). I expected perceived depth to decrease as the contrast of the chromatic component increased.

Figure 5.7: Results from condition 3. Perceived depth as a function of colour contrast for all observers (black circles: sinusoids, grey squares: square-wave, error bars: standard error of the mean, SEM). When raw data showed a significant effect of colour contrast the data was fitted with a Weibull function, the individual fits are presented with solid lines.

Figure 5.7 shows the results for all participants (and Table 5.5 summarises statistical analyses). As before, raw data were analysed with a one-way Kruskal-Wallis test for each participant and each pattern (sinusoids and square-wave). When significant, data were fitted using a Weibull function, either positive or negative (Equation 5.1 & 5.2) depending on the direction of the effect. Curve fits for significant datasets are shown in Figure 5.7.

Table 5.5: Analysis results of condition 3. One-way Kruskal-Wallis test, for each participant and each pattern (sinusoid and square-wave). The effect (increase or decrease) of chromatic contrast on perceived depth is obtained with a linear regression. n.a.: not applicable, see text

	Sinusoid			Square-wave		
Part. ID	$\chi^2(5, 54)$	p	Effect direction	$\chi^2(5, 54)$	p	Effect direction
P1	28.08	p <.001	decrease	18.02	p <.01	decrease
P2	50.32	p <.001	decrease	15.59	p <.01	decrease
P3	38.29	p <.001	decrease	24.34	p <.001	decrease
P4	16.19	p <.01	increase	1.44	.920	*n.s.*
P5	6.51	.259	*n.s.*	n.a.	n.a.	n.a.
P6	28.68	p <.001	increase	1.72	.887	*n.s.*
P7	32.34	p <.001	increase	20.28	p <.01	increase
P8	37.91	p <.001	decrease	11.81	p <.05	increase
P9	5.85	.321	*n.s.*	3.36	.645	*n.s.*
P10	11.32	p <.05	increase	20.56	p <.001	increase
P11	42.18	p <.001	decrease	6.10	.296	*n.s.*
P12	43.24	p <.001	decrease	29.21	p <.001	decrease

As with condition 2, it was found that some participants showed no significant effects of chromatic contrast with depth perception in shape from shading. However, only three participants showed such behaviour in one pattern or both. Participants P1, P2, P3, and P12 showed a decreased perception of depth with increased colour contrast for both sine-wave and square wave stimuli, consistent with previous literature (Kingdom, 2003; Kingdom et al., 2005, 2006). This shows consistency between patterns. However, some participants showed the opposite effect, with a significant increase in perceived depth for both patterns (P7, P10).

For the sinusoids, 10 out of the 12 participants showed a significant effect of colour contrast. Six of them were negative, as predicted by colour shading, the rest positive. For the square wave pattern, only 7 people showed significant effects of colour contrast. Participant 9 showed no significant effect for either the sinusoid or square wave condition. Regarding the step-like results in condition 1, participant P9 seems to have an odd pattern of response and it is possible that this participant was not performing the task properly or using a response strategy.

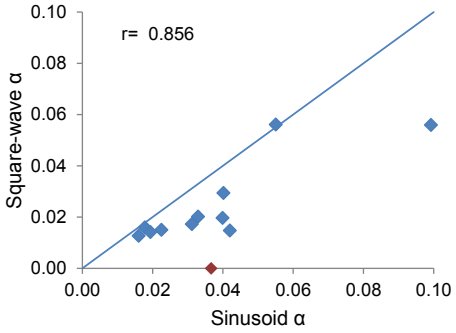
Participant P8 shows a reversal in effect direction between the two stimulus types, delivering a large negative effect for sinusoids wave and a small positive one for the square wave pattern. In total 8 participants showed similar behaviours between sinusoids and square stimulus types. Once again, there was a similar overall pattern of results as for condition 2: a few participants showed no effect of colour contrast on the perceived depth, and some of those with significant effects show a significant increase in the perceived depth in the opposite direction to that expected.

Table 5.6: Fit results for condition 3, all conditions with significant effect (see previous table) were fitted with a cumulative (positive) or inverse Weibull (negative), see Equation 5.1 and 5.2 for details.

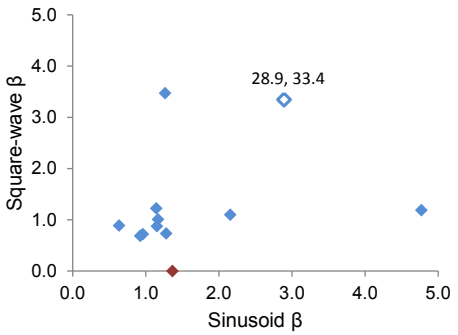
	Sinusoid					Square-wave				
Part. ID	Effect	α	β	γ	σ	Effect	α	β	γ	σ
P1	decrease	0.005	1.260	0.036	1.202	decrease	0.007	1.644	0.281	1.121
P2	decrease	0.008	20.520	0.046	2.005	decrease	0.008	19.886	1.579	1.196
P3	decrease	0.011	2.928	0.284	1.167	decrease	0.007	1.933	0.500	1.064
P4	increase	0.015	60.444	2.607	0.802	no fit				
P5	no fit					*n.a.*				
P6	increase	0.013	1.600	3.171	1.198	no fit				
P7	increase	0.012	2.171	0.683	1.525	increase	0.013	2.616	1.251	0.887
P8	decrease	0.010	4.166	0.162	3.072	increase	0.008	22.955	3.927	0.338
P9	no fit					no fit				
P10	increase	0.004	4.916	3.032	0.701	increase	0.004	14.189	2.888	0.883
P11	decrease	0.009	5.078	0.262	1.935	no fit				
P12	decrease	0.002	0.689	0.000	2.764	decrease	0.0003	0.237	0.000	2.454

Square-wave versus sinusoid

The similarity between the two patterns was apparent in condition 1 (see Figure 5.8). It is evident from Figure 5.8b that square-wave pattern tends to saturate faster than sinusoids.



(a) α parameter, comparison between sinusoid and square-wave. Person r correlation is displayed. The red data point is participant p5 (missing square-wave).



(b) β parameter, comparison between sinusoid and square-wave. Pearson r correlation is displayed. The red data point is participant p5 (missing square-wave). Participant p9 had high values because of step like responses consequently, this participant values are presented with an open symbol and have been rescaled by 10 to fit on the graph.

Figure 5.8: Parameters analysis between sinusoid and square-wave for condition 1.

For condition 2 and 3, when participants had significant effects in both patterns they tended to have the effect going into the same direction (with only one exception in condi-

tion 3). Furthermore, for those with similar effect direction, the data tended to be similar between the two patterns, showing a certain constancy.

In our primary analysis, I did not test sinusoids and square-wave in the same statistical test. This was done because of possible differences in actual contrast of these patterns. I will discuss this point further in the Discussion section.

5.2.3 Discussion

I measured the depth perceived due to shape from shading in a pair of luminance patterns, defined by sinusoids or square waves. The aim was to explore how different combinations of chromatic and luminance components affect perceived depth, i.e. to test the effect of colour contrast on shape from shading. I compared the effects on shape-from shading for both sinusoidal and square wave patterns. Overall, I found a wide variety of observer responses, not all compatible with the heuristic suggested to explain the colour-shading effect, namely that when colour and luminance variations are orthogonal, luminance variation is more likely to be due to shape changes, and hence more depth may be perceived.

In the following sections, I will discuss in detail what was learnt from this experiment. I then suggest that some of our effects, and those of previous studies, could potentially be explained via a hypothesis that is independent of depth perception, using arguments about contrast-masking.

Condition 1: varying luminance contrast for constant colour contrast

I will start by discussing results from a baseline condition (Condition 1) that simply tested how depth from shading increased with luminance contrast. For all participants, the 0 contrast condition always corresponded to a mean perceived depth of 0 arcmin. This result provides evidence that the participants were correctly responding to the shape from shading and that the colour alone did not elicit any depth. For all participants, luminance contrast had a robust effect on perceived depth: increasing the luminance contrast consistently increased the perceived depth. Observers behaved similarly to each other in this condition for both sinusoids and square-wave.

The neutral nature of our matching stimulus, in terms of shape, allowed me to compare the perceived depth directly between the two different stimulus types used here. It was observed that the perceived depth saturates with luminance contrast (Figure 5.5), the speed of saturation varies with the pattern, square-waves tending to saturate faster. However, the maximum perceived depth for both stimuli showed a high correlation (r=0.76) and is similar for the two stimulus types for almost all observers (with the exception of participant P4). This suggests that the perceived depth might not be linked to the geometry of the object depicted by shape-from-shading but instead by the luminance contrast. As sinusoids and square-waves would be expected to have different perceived contrast (Ginsburg, Cannon, & Nelson, 1980) there should be an advantage toward square-waves (more contrast), and this is what was found with square-wave stimuli yielding lower threshold

values (faster saturation, Figure 5.5).

Conditions 2 and 3: the effects of a chromatic grating on perceived depth from luminance

Previous work in this area (Kingdom, 2003; Kingdom et al., 2005, 2006) has suggested that colour enhances perceived depth when presented orthogonally to luminance, and suppresses perceived depth when aligned. In my experiment, the chromatic components seemed to have dramatically different effects on different participants. Some showed the effects expected from the previous literature, some no effects, or some the opposite of the expected behaviour. In condition 2, there was a surprisingly high number of participants with non-significant effects of colour contrast (5.3). For condition 3, colour contrast was supposed to cause less perceived depth. This only occurred consistently for 3 participants (5.5). I will discuss each condition in more detail below.

Condition 2: orthogonal colour component, effects of increasing colour contrast

There are two surprising results for this condition. First, not all participants showed significant effects of colour contrast on perceived depth (see Figure 5.6 and Table 5.3). Participants 2, 3, 8 and 11 showed no significant effects for either sinusoids-waves or square-waves. Most participants produced significant effects for one stimulus type only (P4, P5, P6, P7, P9, P10) plus participants P1 and P12 who whose data shows similar shape in both condition but lack significance for the square-wave. For these people, one could hypothesise that the value of luminance at which they were tested was too small for an effect of colour to be present. This would mean that the chromatic component might interact with shape-from-shading, but only over a specific range of luminance contrasts, specific to each individual.

Condition 1 (see Figure 5.5 and Table 5.1) showed that the depth perceived is directly dependent on the luminance contrast and that the perceived depth saturates more quickly for the square-wave. One could therefore further hypothesize that the perceived contrast is actually different between the sinusoid and square wave. As discussed above, the faster saturation for square wave stimuli in condition 1 suggests that this might be true. In this case, the perceived contrast of the chromatic component might be different as well. Thus, individual differences in perceived contrast could account for the large individual differences in perceived depths found. Overall, these results suggest that if a colour-shading enhancing effect exists, it only does so for a limited contrast range.

The second surprising result is the reduction of perceived depth with chromatic contrast, which was found for some participants. In previous work (Kingdom, 2003; Kingdom et al., 2005, 2006), suppression of perceived depth has only been observed when the chromatic and luminance components are aligned and in phase. These data therefore speaks against a hypothesis based on a *common* set of heuristics suggested previously (Kingdom, 2003) to explain the colour-shading effect.

Could stimulus specific differences explain why my results are not consistent with previous literature? The colour shading effect might be dependent on the luminance contrast at which it is tested. Kingdom et al. (2005) showed a tuning of the colour shading effect depending on the luminance contrast tested, suggesting that there might be smaller enhancement at high luminance contrast. One hypothesis for why in my results were mixed is that the luminance values our observers were tested at might have been too high. If so, perceived depth had already saturated at its maximum level, leaving no space for further enhancement. The luminance contrast used (for conditions 2 and 3) was 4%. (Kingdom et al., 2005) found enhancement for luminance contrasts of 5% and 15%, but less so for 45%.

To address this further, I considered all the individual data and looked for correspondence between the luminance level tested in the chromatic conditions (i.e. Figures 5.6 & 5.7) and the level of saturation of contrast in condition 1. If the above hypothesis were correct, then I might expect to see non-significant effects or suppressive effects at saturation levels of luminance, and increasing facilitatory effects for those observers who were not yet at their saturation level. However, this pattern was not found consistently, some participants showed increases or decreases at saturated value and others non-significant effects, even though they were tested at their mid-range. Thus, the saturation hypothesis does not explain why some participants showed non-significant effects of colour contrast.

An additional difference is the fact that I used only two components in the design of this experiment and not 3 as in most previous experiments. However, if a common heuristic is used to perform shape versus reflectance segregation then it should apply to any stimulus. If the effect is stimulus specific then it might not be linked to what was hypothesized.

Condition 3: luminance and colour contrast components aligned

When luminance and colour signals were aligned, I expected decreased perceived depth, consistent with a suppression of luminance-defined depth at borders that specify both luminance and chromatic changes. This was not consistently found. Compared to condition 2, more participants showed significant effects of colour contrast when the two components were aligned (compare Figures 5.6 & 5.7). There was greater depth modulation as a function of colour contrast for the sinusoidal pattern than for the square-wave pattern (see 5.5, and 5.7). Note also that the direction of the effects, when present, tended to be the same for sinusoids and square waves. However, a few participants showed non-significant effects of colour contrast increase on perceived depth.

Some participants with significant effects showed the opposite of the expected effect direction, i.e. they showed an increase in perceived depth. One example is participant 7 (see Figure 5.7). Thus, once again, I do not find strong evidence for a systematic effect of colour that would be consistent with the heuristic suggested to explain the colour shading effect. Furthermore, in the previous section I argued that some decreases in perceived

depth might be because the luminance contrast was too high (i.e. the maximum depth was already reached), but this argument does not match the observed data in condition 3. Participant P4 exemplifies this assessment. Consider their behaviour for condition 1 (see Figure 5.5) and compare with 2 and 3 (see Figures 5.6 & 5.7). The values of luminance contrast tested in conditions 2 and 3 are already in the saturated part of the curve for the sinusoid, and a suppression of perceived depth is in effect when colour is orthogonal (Figure 5.6) and an increase when aligned (see Figure 5.7), both significant see Tables 5.3 & 5.5. This is the opposite from what one would expect if perceived contrast were the only issue.

Other stimulus details differed between my study and previous ones. A key difference between this experiment and those performed in other labs was that in my case the depth suppressing effect was not tested using two chromatic components: one iso- and one cross-oriented, but only one iso-oriented component. Kingdom et al. (2005) showed that a luminance component would have its depth enhanced by an orthogonal chromatic component (either L-M or S modulated) and suppressed by a third colour component (either L-M or S modulated) aligned with the luminance component. In condition 3, I used only one (aligned) colour component. According to the colour-shading effect hypothesis, the alignment of chromatic and luminance should be interpreted as changes in reflectance, and hence perceived depth should be reduced. This kind of interactions found in previous studies could also be attributed to interactions between the two chromatic components. I explore this idea in more details in Section 5.4.4.

Luminance artefact

In any colour vision experiment, there is the possibility that chromatic information has not been accurately isolated and that small luminance artefacts could underlie some of the observed behaviour. In this section, I discuss several reasons why luminance artefacts are very unlikely account for our results.

Firstly, chromatic aberration can occur due to the differential refraction of the eye however it is only a problem at medium and high spatial frequencies (Mullen, 1985). For the spatial frequencies used here, (0.75 cpd) this should not be an issue.

The second possibility is that despite having tested individual isoluminant points for each participant and contrast value tested, there still exists some luminance artefact. The results of condition 1, that showed how perceived depth increased with contrast for every observer (Figure 5.5), provides a compelling case that the depth perceived in our stimulus was mainly driven by luminance contrast. Any manipulation that decreases the perceived contrast of the right oblique luminance component should then also affect the perceived depth. Georgeson and Shackleton (1994) found, for a detection task that plaids have less perceived contrast than individual gratings, and that it was minimal for orthogonal gratings. If there were a luminance artefact at work in our condition 2 (luminance and chromatic gratings orthogonal to each other), then it should decrease the perceive contrast

of the shape-from-shading object. This argument was first formulated by Kingdom et al. (2005). This might be a factor for those of our participants that showed a decrease in perceived depth with colour contrast (P4, P6 and P9). However, for the remaining 9 participants, where no change or an increase in perceived depth was found, this is not borne out.

Furthermore, in condition 3, where chromatic and luminance components are aligned, any luminance artefact in the chromatic component should directly add to the luminance one and result in enhanced depth or suppressed depth, depending on the orientation.

In order to explore the possibility of luminance artefact further I re-tested some participants with modified stimulus phase, as described in the second part of this chapter.

5.3 Part 2

5.3.1 Aims

The experiment was aimed at testing the possibility of luminance artefact in the stimuli. Luminance artifacts are to be avoided for obvious reasons when using stimuli based and shape-from shading, as the actual luminance shading would be distorted.

I describe our equiluminance settings in a separated Chapter 8, the results of the equiluminance experiment are also presented for every participant (for each contrast values tested and each pattern type).

Because I did not counter-balance the phase of the stimuli in condition 3 (the luminance and chromatic phase had always the same phase relationship), it is possible to check for luminance artefact as described in the following section.

5.3.2 Method

The artefact retest experiment is based on the alignment of luminance and chromatic (red-green) components in the third condition. By changing the phase, it is possible to detect irregularities if participants change behaviours. As the same spatial frequency was used for both stimuli (see Method section), any luminance component in the chromatic stimulus would be additively combined with the luminance defined frames. Therefore, this retest is based on the results of condition 3, if the behaviour of the participants change on this condition then there would be strong evidence for luminance artefact. If there were no change, I would have a stronger argument against the presence of artefact however, absence of proof is not sufficient to claim they do not exist. One could argue that our protocol does not capture luminance artefact properly or that it is not sensitive enough for small artefacts, these issues will be discussed in Section 5.3.4 after the results.

Observers

As the test relies on results of condition 3, I selected two participants with at least one significant result in the third condition (5.5). Based on their response on the third condition, i.e. I selected one participant with a decrease in perceived depth (participant P7) and one with a decrease (participant P11).

Stimuli and conditions

The stimuli used were similar than described in Part 1 (see Section 5.2.1), the only difference was the relative phase of luminance and chromatic component, this was obtained by changing the definition of the sinusoids to add a phase offset of $+180^\circ$ or π (in radians) to the chromatic component.

I retested both participant on both sinusoids and square-wave experiment the participant did not re-perform the equiluminant experiment and I used the data from their previous experiment in order to have the same stimuli amplitude but different phase. As

the test and re-test were performed only few days apart I did not expect major drift in isoluminant point (this point is addressed in the Equiluminance Chapter 8, see data on test retest consistency, also see Data comparing sinusoid versus square-wave).

5.3.3 Results

In this section I present the results of condition 3 that answer the main question but also the results of condition 1 & 2 which gives a good look at the repeatability of the results obtained previously.

Condition 3-restest: main result

The results are presented in Figure 5.9. The test and retest values are presented on the same graph. The retest follows the same direction than the original data. It is especially striking for participant P11 who had only a significant effect in the sinusoid condition and not in the square-wave.

These results give more confidence towards the absence of luminance artefacts. Alternatively, if there are still such artefacts present we conclude they have no effects on our task and on the observer perception.

Additionally, participant P7 tends to have similar standard errors between test and retest whereas P11 shows an increase in variability.

Condition 1-retest

As stated earlier these results (i.e. condition 1, orthogonal luminance contrast) cannot answer the question about luminance artefact. However, as re-test experiment was necessary to check for artefact (with analysis on condition 3) there is additional data on the robustness of the effects shown in Part I.

The results are presented in Figure 5.10, this figure uses the same structure than in Figure 5.9, it presents both original and retest values, for both participants (top and bottom row) segregated by pattern (left and right column).

In both participants' data, it is clear that the general shape of the data is similar; however, some scaling seems to be present between sessions. This is especially the case for participant P11. Remember that no differences were expected in this condition as the luminance pattern is neither shifted nor changed in value.

Participant P7 shows less so this effect and in the opposite direction than participant P11. This is important as this might give us pointers towards the matching task per se and probably less so enlightening the shape from shading.

Note that when going back to the results in condition 3, the same effect (see Figure 5.9) seems visible, with a slightly lower value for p7 between test and retest (although these are only non-significant numerical differences). Participant p11 shows higher values in condition 1 when no chromatic component is present.

Figure 5.9: Results of retest: condition 3. The original data (from Part I) is presented (filled symbols) with the retest phase shifted data (open symbols). Participant p7 data is on the top row, and participant p11 on the bottom row. Sinusoid data is presented on the left column (circle symbols) and square-wave data in the right column (square symbols). Error bars: SEM. The solid line represents fit from original data when applicable. No reversal of effect was found.

Condition 2-retest

Similarly to the previous section the results, what is described here does not address directly the luminance artefact question but the regularity of data across time from participants. The results are presented in Figure 5.11, this figure uses the same structure than in Figures 5.9 & 5.10.

In both participants' data, the general shape of the data is similar; however, some scaling seems to be present between sessions. This is especially the case for participant P11. Remember that there should be no differences in this condition as the luminance pattern is neither shifted nor changed in value.

From the two participants, only p7 showed a significant effect in the orthogonal condition (chromatic contrast as independent variable). As from the previous section, participant p7 has generally lower values in the retest (open symbols). This scaling seems to reduce the increase effect found earlier. Likewise, in condition 3, for participant P11, the retest values are consistently higher than the original.

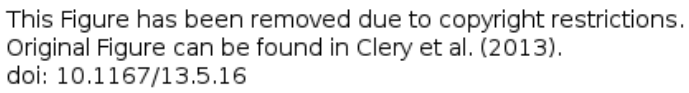


Figure 5.10: Results of retest: condition 1, orthogonal luminance changes. In this condition, the chromatic phase is irrelevant; hence, there should be no major differences. The original data (from Part I) is presented (filled symbols) with the retest phase shifted data (open symbols). Participant p7 data is on the top row, and participant p11 on the bottom row. Sinusoid data is presented on the left column (and circle symbols) and square-wave data in the right column (and square symbols). Error bars: SEM. The solid line represents fit from original data. Participant p11 seems to show a scaling effect between test and re-test.

5.3.4 Discussion

Lack of evidence for luminance artefacts

There are several points that need to be discussed regarding the results and their potential interpretations. I first briefly summarize our initial argument why this experiment is useful for artefacts search and I then discuss the limitations, experimental and dialectic regarding the interpretation.

There were no major differences in direction or strength (taking into account the scaling differences) between test and re-test. The irreducible fact that can be derived from this is that chromatic stripes (red and green) have the same effects on perception of depth from shading modulation whether they are aligned with dark or bright stripes of the luminance stimulus (for this two participants).

Now the basic premise can be interpreted further following the hypothesis on artefacts.

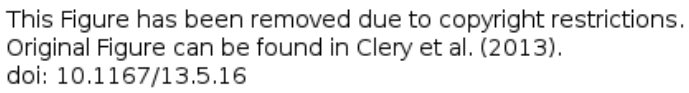


Figure 5.11: Results of retest: condition 2. The original data (from Part I) is presented (filled symbols) with the retest phase shifted data (open symbols). Participant p7 data is on the top row, and participant p11 on the bottom row. Sinusoid data is presented on the left column (and circle symbols) and square-wave data in the right column (and square symbols). Error bars: SEM. The solid line represents fit from original data when applicable.

If a consistent artefact was present, consistently across all contrast values (with the exception of null contrast condition), then by shifting the artefact effect onto the luminance bands the effect off the artefact should be reversed. This makes the additional assumption that chromatic bands (red or green) interact in a similar way with luminance (and there is no evidence of the opposite in the previous literature on colour shading effect).

I conclude that probably no artefact is responsible for the effect seen in condition 3 for these two participants. Two extensions can be made, one regarding the results in orthogonal colour effects and one regarding others participants. However, our test limitations cannot provide definite answers.

Shape from shading: Effect robustness

The experiment in Part II gave us information about the reliability of the results over time for two of our participants. The directions of effects were preserved in condition 3. Figure 5.9 shows that the fit from the initial data (when present) matches the shape of

the re-test data.

The condition 1 re-test indicates that there might be a scaling between the two sessions; however, the overall shape seems to be preserved too. Assuming that the perception of shape from shading is a robust effect that does not vary over time then the scaling might happen when the participants transfer the shape perceived in one dimension to another.

The task is not straightforward, as the two percepts have different qualia. There is some evidence that the two visual information, from shading and stereopsis, are combined (Bülthoff & Mallot, 1988); and that it is done optimally (Lovell et al., 2012). There is some evidence that the combination is fast and dynamical regarding to how noisy the channels are (Lovell et al., submitted). Furthermore, the shape from shading stimulus is providing a conflicting stimulus as the stereopsis information is flat but the shape from shading is indicating shape variations. This is a difficult task as can be seen with participant p5 who could not perceive shape from shading reliably with the square-wave stimulus.

Despite the difficulty, the re-test in condition 1 seems to show that our participant can perform the task reliably. I hypothesize that the scaling factor (more evident in participant P11 than P7) refers to the depth estimation mapping between two depth modalities.

Alternatively, one could assume that the mapping is fixed but that the overall shape from shading depth perceived has increased. This would be due to learning and increase familiarity with the stimuli and task. Dovencioglu, Schofield, and Welchman (2011) have shown that Second order luminance modulation can be learned dynamically very quickly. I will come back to this point in the general discussion within the context of chromatic modulations.

5.4 General discussion

5.4.1 Summary

In summary, firstly I showed that the participants were able to perform the task in two separate sessions (sinusoid and square wave) with consistency. The sinusoid and square-wave pattern evoke different patterns of shapes (P. Sun & Schofield, 2012), however their relation to luminance contrast is very similar. The effect of contrast in dependent on the participant in term of shape and threshold but within participant there is mostly high resemblance between sinusoid and square-wave processing. I note however, that participant p4 had scaling (or magnitude) differences between the two patterns but the shape and threshold were nevertheless similar. Secondly, one of our participants (P5) could not see any consistent depth with the square-wave pattern. I did not exclude this participant because this confirms data by P. Sun and Schofield (2012). In their paper P. Sun and Schofield (2012) did not search explicitly for contrast effects however they show data for three participants in which one can compare sinusoid versus square wave. The profile depth perceived seems to match between sine and square-wave but also some

participants could not see depth in the square-wave modulations.[3]

5.4.2 Choice of task and participants' strategy

In the original study on the colour-shading effect, the matching stimulus used to assess the shape of the stimuli was a stereo-grating (100% contrast stripes) whose shape was modulated to have a sinusoidal 3D shape when fused. The depth corrugation was tilted to the left, the same orientation than the luminance stimuli to be matched. The size of the corrugations was of the same spatial frequency as the colour-shading stimuli. The task was to match the depth of the corrugations in the colour-shading stimuli (luminance and colour variations) with the matching stimuli (stereoscopic depth corrugations). Kingdom et al. (2005) used a modified version of the matching stimuli. The stereo was delivered using random dot stereograms (as opposed to vertical stripes). The depth profile was composed of a sinusoidal modulation oriented to the left, as in Kingdom (2003). Although the authors noted that they discovered in a pilot experiment that the percept of shape from the sinusoid stereogram did not match the depth from shape-from-shading of a luminance sinusoid, the shape-from-shading percept was smoother. In order to compensate, Kingdom et al. (2005) added harmonics to the main spatial frequency that rounded the percept of shape. As in Kingdom (2003), the amplitude of the pattern from highest to lowest point was used as a measure of stimulus corrugation. Finally, Kingdom et al. (2006) used a similar matching stimulus as in Kingdom et al. (2005) as well as a simple matching bar protocol. The matching bar was a bar that could be adjusted in length on the screen; the observers reported the magnitude of the depth observed with this bar serving as a ruler. By comparing both techniques, the authors did find that the pattern of answers was similar; participants were able to extract the peak to trough depth from one dimension and transfer it to another one.

There are two important points regarding this stimulus. Firstly, the reconstruction of shape-from-shading needs to be taken into account. Importantly, a sinusoid luminance variation does not translate to a pure sinusoidal shape. P. Sun and Schofield (2012) studied the perceived slant of sinusoid and square-wave stimuli. They demonstrated differences between participants and non-regularities in the perception of square-waves. It can therefore be problematic to create appropriate matching stimuli (e.g. Kingdom et al., 2005). In the case of my experiment, the use of square-wave patterns would have required to make assumptions about the geometry of the final percept of each observer. Secondly, Kingdom et al. (2006) demonstrated the ability of participants to estimate depth amplitude in one modality and transfer it to another. These magnitude estimations are used extensively in psychology experiments; however, we do not yet understand how this process works.

In conclusion, regarding the choice of matching stimuli, I believe my matching stimulus was a middle ground between a realistic shape matching stimulus (which would require assumptions about shape) and a purely abstract bar match (which does not convey the

[3]This is dependent on the orientation of the profile.

same depth percept as my neutral depth match stimulus). In my modified matching task, there are no assumptions at all about the shape of the object. It is possible for the participant to adopt a strategy to assess the depth within the stimuli by trying to find where the closest plane (peak) and the other plane (trough) of the stimuli lie. If the task is performed in this manner, this should not influence the interpretation of the results.

However, it is possible that participants switch to a (contrast) magnitude estimation task whenever the shape percept is unclear or ambiguous. The process of magnitude estimation can in itself be the source of individual differences between participants (Green & Luce, 1974). This could explain some of the apparent differences in shape perception within my sample.

5.4.3 Link with second order variations

The literature on second order luminance effects is also of bearing here. A first order variation is luminance change across space, what is usually described as contrast. A second order variation is a change of contrast across space. The second order luminance modulation is relevant to the effect of chromatic information on luminance-defined-shading, for two main reasons. First, because the channels responsible for detection of LM (local mean luminance, first order) and AM (local luminance amplitude modulation), sometimes also referred as contrast modulation, CM (Schofield et al., 2006), interact in a similar way to luminance and (red-green) chromatic channels (see Schofield & Georgeson, 1999; Chen et al., 2000a). Secondly, it has been shown that both these channels have modulatory effects on shape-from-shading processing (for second order luminance: Schofield et al., 2006; and for chromatic components see Kingdom, 2003). Note that this connection has also been made by Schofield et al. (2006) and by Kingdom (2008).

One particular experiment by Dovencioglu et al. (2011) showed that AM-LM interactions could be learned over time by naïve observers using positive reinforcement. From this work, two conclusions seem important. First, naïve participants might not be using these interactions unless trained. Therefore, it is a possibility that luminance and colour interaction could be learned as well. Second, it indirectly raises a cautionary note on the use of trained observers or non-naive (usually authors) trained in psychophysics. Our relatively modest sample size of 12 naïve participants has shown tremendous inter-individual differences in the effects of colour contrast. However, the effects of luminance contrast on perceived depth were similar across all participants and robust between sinusoids and square-waves. It has been suggested, in the case of a different mechanism: motion-in-depth that the behaviour of the general population can be quite be different that of trained 'expert observers' (Nefs, O'Hare, & Harris, 2010), and this is very much in line with what was found here.

5.4.4 Masking hypothesis

One alternative way to interpret our results is to think in terms of masking. In other words, to what extent chromatic patterns might mask luminance patterns, at stages well before the luminance information is interpreted as depth. Most low-level interactions between sinusoidal luminance, or colour, patterns are described in term of masking within channels (e.g. Legge & Foley, 1980; Losada & Mullen, 1994) and between channels (e.g. Chen et al., 2000a).

Masking experiments consist of testing the effect of an additional stimulus (mask) added to a target stimulus that the observer is asked to detect or discriminate (Legge & Foley, 1980). In the spatial vision literature, sinusoidal modulations of luminance are commonly used for mask and signal. Historically, this technique was developed to study interference within the same channel, i.e. the effects of luminance masks on luminance tests, or colour masks on colour tests. Many results from the previous literature are therefore not directly relevant to the work presented here. Further, masking experiments are typically conducted at, or near, detection thresholds and our stimuli were clearly suprathreshold. However, it is still of value to consider these studies and explore whether effects in our own work could correspond to those found in masking studies.

An interesting effect in masking is the facilitation that occurs when the mask is at detection level. The facilitation (manifested as a lowered threshold) is then followed by an increase of the threshold at higher contrast. The resulting pattern threshold elevation, as a function of mask contrast, is commonly referred to as the 'dipper function' because of this initial facilitation, then suppression effect. This pattern of masking occurs within chromatic (e.g. Losada & Mullen, 1994) as well as within luminance channels, e.g. Bird et al., 2002).

The strength of the mask is dependent on the characteristics of both mask and target. As a rule of thumb, the more similar mask and target are, the greater the masking effect. This is explained by hypothesising that the same channel processes test and mask, and that the mask is adding noise, rendering detection more difficult. For example, masking of a sinusoidal luminance pattern is greatest when the mask and test are about the same orientation (Phillips & Wilson, 1984), but cross-orientation masking can also occur: targets are still masked when masks are 90° out of phase (Meese & Holmes, 2007; Phillips & Wilson, 1984).

In colour vision, masking follows a different pattern: masking appears to be isotropic, i.e. not dependent on mask orientation (Medina & Mullen, 2009). An orthogonally oriented (also referred to as cross-oriented) stimulus usually shows suppressive effects, but masking can be facilitatory (Meese & Holmes, 2007) at low temporal frequency. Medina and Mullen (2009) compared cross orientation masking (XOM) within luminance and chromatic channels. Chromatic XOM was found to be independent of temporal frequency and generally stronger than achromatic XOM, i.e. stronger facilitation at low contrast and stronger suppressive masking at higher contrast.

Such effects could provide a potential link to explaining the differences between our results and others. Two components were used per stimulus, one achromatic and one chromatic. Thus, there will be no within-channel masking. XOM, i.e. when two chromatic components are present and orthogonal to each other, might be responsible for the suppressive effects found for the three-component stimuli used in Kingdoms studies.

Studies on how colour masks luminance or vice versa (cross-channel masking), are rare (but see Gheiratmand & Mullen, 2010, Mullen, Gheiratmand, Medina, & Kim, 2012, Chen et al., 2000a) and typically conducted at, or near, detection thresholds. Hence, one cannot yet predict from them whether specific aspects of the colour shading effect could be accounted for by low-level masking. There is also some data on cross-channel (red-green chromatic vs. luminance) iso-oriented interaction. Chen et al. (2000a) found that chromatic masks interact, but do not facilitate, luminance target detection when the chromatic component is properly isoluminant (see section on luminance artefacts below), however when it is not they found slight facilitatory effects at low contrast. Iso-oriented masking could potentially be involved in the suppressive effect that was found for some participants (Figure 5.7) and others have found (Kingdom, 2003; Kingdom et al., 2005).

In the supra-threshold domain, Pearson and Kingdom (2002) found that superimposed chromatic (red-green) and luminance gabors showed suprathreshold facilitation, as opposed to the sub-threshold facilitation showed in detection tasks (dipper function behaviour). This study also showed marked individual differences in stimulus integration (for example, one participant did not show any facilitation, but only the suppressive effect, with a luminance target and chromatic mask). From the results of our condition 3, we can hypothesize that some participants with increased perceived depth might misappropriate chromatic contrast for luminance contrast; this supra-threshold combination of contrast might be akin to results described in Pearson and Kingdom (2002). It is possible that the aligned configuration is perceived as ambiguous, one hypothesis could be that the respective contrasts (luminance and colour) are pooled together, or colour contrast somehow modulates luminance contrast. If this pooling mechanism exists, it could well occur before the shape is computed.

Gowdy et al. (1999) tested XOM with sine and square wave patterns. They found that at 0.8 cpd there was larger facilitation for sinusoids than square-waves. They concluded that the sharp-edges played an important role in the facilitation. They also found that the facilitation was phase dependent. The facilitation was similar at 0° or 180° phase (green on light bar or green bar on dark bar) however at 90° the facilitation was abolished. This is interesting when attempting to interpret the colour-shading effect because we know from Kingdom (2003) that the colour-shading effect is also phase dependent. However, DeValois and Switkes (1983) found suppressive (masking) effects of chromatic masks on luminance targets. Two of their participants showed no phase-effects but one did show a large suppressive effect for 0°and 180° (in-phase) and a small effect for 90°. In their paper, they postulated that luminance and chromatic masks could have distinct excitatory or inhibitory effects on the detection of luminance targets.

To sum up, it is difficult to know how the masking data obtained from other studies relates to the perceived supra-threshold contrast that was used in our study and in others on the colour-shading effect. However, the phase dependency and orientation seems to be similar in masking and suprathreshold data; both can show large individual differences too. It would therefore be interesting, but beyond the scope of this study, to measure how the different chromatic components affects the perceived contrast of the luminance modulation.

5.4.5 Conclusions

I have shown that sinusoidal and square-waves luminance modulations have similar effects on perceived depth. I interpreted the faster increase of perceived depth with contrast of the square-wave as a sign of higher perceived contrast between sinusoids and square-waves. However, the maximum perceived depth was similar for both patterns for 11 of our 12 participants.

For the chromatic contrast manipulation, both iso- and cross-oriented chromatic contrast increases were perceived with highly individual differences by the test group. Therefore, I do not think that a common heuristic is used by the whole population to disambiguate luminance variation that could be delivered by a shape or a reflectance change. Dovencioglu et al. (2011) suggest that a similar relationship can be learned for AM-LM modulation with feedback. Consequently, I argue that people usually described in the literature as 'expert observers' might have learnt how to use a common heuristic to disambiguate the luminance information.

The interactions that were observed for the iso-oriented (aligned) conditions are akin to masking and an alternate explanation of the colour-shading effect could potentially be accounted for with low-level masking. As I described, the literature contains multiple examples of low-level interaction between channels. The next step in this thesis is to explore the possibility of low-level inter-individual differences in the processing of luminance and chromatic contrast. Furthermore, a complete account of these interactions should test for suprathreshold inter-channel interactions. These approaches are detailed into following data chapters.

Chapter 6

Suprathreshold contrast interactions

6.1 Introduction

The aim of this chapter is to explore the link between suprathreshold contrast processing and shape-from-shading processing. If sequential processing is assumed, i.e. that contrast processing information feeds into shape-from-shading, then any early integration and interaction of luminance and chromatic information should have an effect on shape-from-shading. The formal hypothesis is that low-level interactions are responsible for the effect found in the previous chapter, i.e. modulation of the amplitude of shape perception of luminance patterns by chromatic patterns (sinusoids or square-waves). These effects were variable between participants, as opposed to luminance effects alone, which acted as the carrier for the object shape.

The modulatory effect of colour has been suggested to be due to the visual system using heuristics to perform image (or scene) segmentation of reflectance versus shape (Kingdom, 2003). In the previous chapter, Chapter 5, my results raised doubts on the general validity of the heuristics hypothesis. The results did not deliver enough evidence to revoke the idea, as it is possible that participants used different heuristics (explaining the individual differences). However the fact that some of the variation of effects with contrast went in the other direction than expected from the segregation hypothesis, questions the validity of this part of Kingdom's (2003) hypothesis. I also advanced an alternative hypothesis, where low-level masking caused the effects, without being linked to scene processing specificity.

In this chapter, I will test an alternate explanation for the colour-shading effect presented in the previous chapter. I hypothesize that an interaction between the chromatic and the luminance channel could be the mechanism behind the colour-shading effect. This low-level interaction would modulate the putative luminance signal that is then used for higher-level tasks such as shape-from-shading. Consequently, if chromatic contrast can affect luminance contrast processing, as I expect it will, then the luminance signal can be suppressed or enhanced depending on the configuration and the participant. The ex-

pectation is to find suppressive and enhancing effects tuned to chromatic contrast with variations in the population comparable to the variability in the colour-shading effect found previously.

Furthermore, on the issue of heuristics, it is not clear how heuristics computation is done (see Kingdom, 2008, for an extensive list of heuristics in shape processing), and it is possible that low-level interactions and high-level feedback are both present. It is difficult to disambiguate the source of the effect between heuristics or low-level contrast processing as the two might be intertwined. It is possible to treat this issue using Marr's (1982) approach, if we consider the problem to be that of solving shape extraction and the underlying computational solution involving contrast manipulation. It is possible to gain understanding of this problem by accessing the lower layers of processing. A marker of such low-level interaction is contrast perception itself. That is why I will focus on suprathreshold contrast perception in this chapter and specifically on how luminance contrast perception is altered by the presence of colour stimuli.

The experiment presented in this chapter focuses on orthogonal interactions; it aims at comparing the orthogonal effects of luminance and colour on a luminance target. This is presented in Section 6.2. I then compare the results obtained from this experiment with the previous data on colour shading for participants having performed both experiments to check if my hypothesis that low level interactions are responsible for the colour-shading effect is valid (Section 6.2.4) and discuss the implications.

There have been suggestions in the literature about the specificity and complexity of suprathreshold perception (Cannon & Fullenkamp, 1993; Switkes & Crognale, 1999; Vimal, 2000). This chapter will explore new aspects of contrast perception and its interactions with other processes.

6.2 Experiment II

6.2.1 Aims

I designed this experiment in order to investigate supra-threshold interactions between red-green chromatic and luminance channels. The focus was on contrast perception of a luminance target and the influence of orthogonal masks.

The aim was to compare the processing of luminance masks and colour masks. From the literature, I expected overlaid orthogonal luminance masks to have a suppressive effect on the perceived contrast of a luminance target (see for example, lateral inhibition, Physiology Chapter 2). Such modulatory effects are expected to be tuned depending on the contrast of the target and the masks. For this reason, in this experiment, I tested a range of contrasts of both targets and masks (see Section 6.2.2). Furthermore, as this thesis topic is channel interactions, I will also be looking into differences between chromatic (red-green) and luminance masks' modulatory effects on suprathreshold perception of luminance targets.

Consequently, I will be measuring the suprathreshold perception of luminance targets in the presence of luminance or chromatic orthogonal masks of varying contrast.

As explained in the introduction of this chapter, ultimately the goal is to infer the processing stage before shape perception and search for possible low-level interactions responsible for modulatory effects of shape-from-shading, or in a simpler form: how the channels are bound together for higher visual processing.

Here, I am aiming at uncovering some of that processing. Furthermore, I tried to test a large group of participants in order to check for individual variations of this processing in the population. I also tried to re-test participants from the shape-from-shading experiment, in order to match the results of contrast perception and shape-from-shading.

6.2.2 Methods

Observers

A number of 22 observers were tested. All observers were naïve to the task and hypothesis with the exception of the author (M1)[1]. All of them had no colour deficiency as tested by the Ishihara colour plate test (38-plate edition, 1979) and had normal or corrected-to-normal acuity. Ethical approval was given by the University of St. Andrews ethics committee UTREC (reference PS6135) and followed university guidelines. Participants were given monetary compensation for their time.

For this experiment I tried to recruit participants who had also participated in the experiments in Chapter 5. Out of the 12 participants in those experiments, 8 were retested here (the two experiments were performed about a year apart). The details of the participants retested from the previous experimental chapter are provided in Table 6.1.

Table 6.1: IDs of participants tested in both the depth and the contrast matching experiments

Experiment	Participant ID							
Matching experiment	M4	M10	M14	M15	M16	M17	M18	M19
Colour shading exp.	P3	P10	P5	P7	P11	P2	P6	P4

Apparatus

The stimuli were displayed on a CRT, Iiyama Vision Master Pro 22″, (cadenced at 100 Hz, non-interlaced). The CIE-1931 chromaticity coordinates of the red, green, and blue phosphors were $x = 0.623$, $y = 0.342$; $x = 0.289$, $y = 0.608$ and $x = 0.149$, $y = 0.073$, respectively. The gun chromaticity values were obtained by measuring the full spectra

[1]The nomenclature used in this chapter to distinguish the participants from those in Experiment I is M1-M22, the author is M1. This code was used to distinguish participants from the depth matching experiment labelled P1 to P12. When participants performed both experiments both codes are given as a pair, when necessary, e.g. M4/P3 (see Results section 6.2.4)

of the gun with a PR-650 (Photo Research) spectroradiometer and then converting into CIE-1931 (2° Standard observer) values by multiplying the functions and spectra. The stimuli were generated using a ViSaGe system (Cambridge Research Systems Ltd.) and Matlab (Mathworks) running on a Dell workstation. The screen gamma non-linearity was corrected using a CRS ColorCal colorimeter. Details on stimuli are defined in the following sections and details on colour definition in Chapter 4.

Participants viewed the screen dichoptically in a darkened room; a chinrest was mounted to hold the observers head steady at a viewing distance of 96 cm.

Stimuli and task: General

The experiment consisted of a luminance contrast-matching task using the method of adjustment. For each trial, two disks were presented, the upper disk contained the target and the lower disk was the matching stimulus that could be adjusted to match the perceived contrast. In order to check for modulatory effects, the target probe had orthogonal masks of different types and contrasts superimposed (see General Method Chapter 4 on interleaving, Section 4.3.3).

The upper stimulus consisted of target plus mask combinations. The target was a sinusoid in luminance. The mask was either a luminance or chromatically defined sinusoid and always orthogonal to the target. Participants were asked to match the perceived contrast in the left-oblique direction (target) and ignore distractors in the other direction (orthogonal mask). The matching probe was composed of a luminance sinusoid only, always presented in the left oblique direction, the same as the target. Observers had no time constraint to make the match.

All sinusoids had the same spatial frequency (0.75 cpd) as the experiment in Chapter 5. The size of the stimuli was 4°, the edges were smoothed using a Gaussian filter. I described stimulus generation and colour space in a previous chapter, however, I give a brief overview here.

The background intensity used was $x = 0.282$, $y = 0.311$ and $Y = 40\ cd.m^{-2}$. These values were transformed into cone excitation, which were obtained using Brainard et al.'s (2002) transform from xyY CIE to normalized V. C. Smith and Pokorny (1975) cone sensitivity (see Methods Chapter, cone space Section 4.2.6, p. 82).

Equations 4.7, 4.8 & 4.9 define the sinusoidal modulation of contrast; A corresponds to the peak-to-peak amplitude modulation. When I refer to contrast value hereafter, I refer to a from Equation 4.13. For luminance modulation all cone contrast classes are modulated by A and kept in-phase. The value b is only used in chromatic components to adjust for iso-luminance differences between participants. Values of b are dependent on A. Positive b values: L cone contrast is increased and M contrast is decreased, and vice-versa for negative b values.

Table 6.2: Details of conditions tested are presented in mask\target pairs, in contrast (%).

			Target contrast				
			Baseline	Level 1	Level 2	Level 3	Level 4
Masks	Baseline		0\0	0\0.02	0\0.075	0\0.16	0\0.5
	Colour	Low	0.01\0	0.01\0.02	0.01\0.075	0.01\0.16	0.01\0.5
		High	0.03\0	0.03\0.02	0.03\0.075	0.03\0.16	0.03\0.5
	Luminance	Low	0.01\0	0.01\0.02	0.01\0.075	0.01\0.16	0.01\0.5
		Med	0.02\0	0.02\0.02	0.02\0.075	0.02\0.16	0.02\0.5
		High	0.10\0	0.10\0.02	0.10\0.075	0.10\0.16	0.10\0.5

Equiluminant setting

To obtain a measure of b, participants performed hetero-flicker photometry (Anstis & Cavanagh, 1983). The hetero-flicker photometry experiment consists of alternating colour sinusoids, 180° out of phase. When red and green stripes are at apparent equal-luminance, then the flicker perception is minimal. The method of adjustment was used. By turning a response knob (Cambridge Research Systems CB7), each observer adjusted the relative amount of green/red contrast. The full description of the equiluminant adjustment experiment as well as data and additional analysis is presented in the Isoluminance Chapter 8.

Design: Conditions

There were 5 different target contrasts tested ranging from zero contrast (acting as a control) to high contrast values (0, 0.02, 0.075, 0.16 and 0.5).[2] There were 6 orthogonal mask conditions: 2 chromatic contrasts (0.01 and 0.03, referred to as low/high), 3 luminance contrasts (0.01, 0.02, 0.10, referred to as low, medium, high) and a no mask condition (baseline). Each target contrast was tested against each mask contrast resulting in 30 conditions. All conditions are shown in Table 6.2 and are displayed in Figure 6.1.

Design: Sessions

Each session comprised 10 repeats of all conditions. There were 30 conditions in total, so each session had 300 randomized trials. Participants performed a flicker photometry test before the first session to calibrate the experiment. Each participant performed four sessions approximately lasting one hour each. At the end, participants were tested once again on the flicker photometry to check for drifts and validate the calibration (see Equiluminance Chapter 8, Section 8.4.3, p. 215). At the end of the 4 sessions, the dataset consisted of 40 trials per condition.

[2]Units are modulations of cone contrast excitation. Parameter A multiplies the sinusoids of cone contrast modulation.

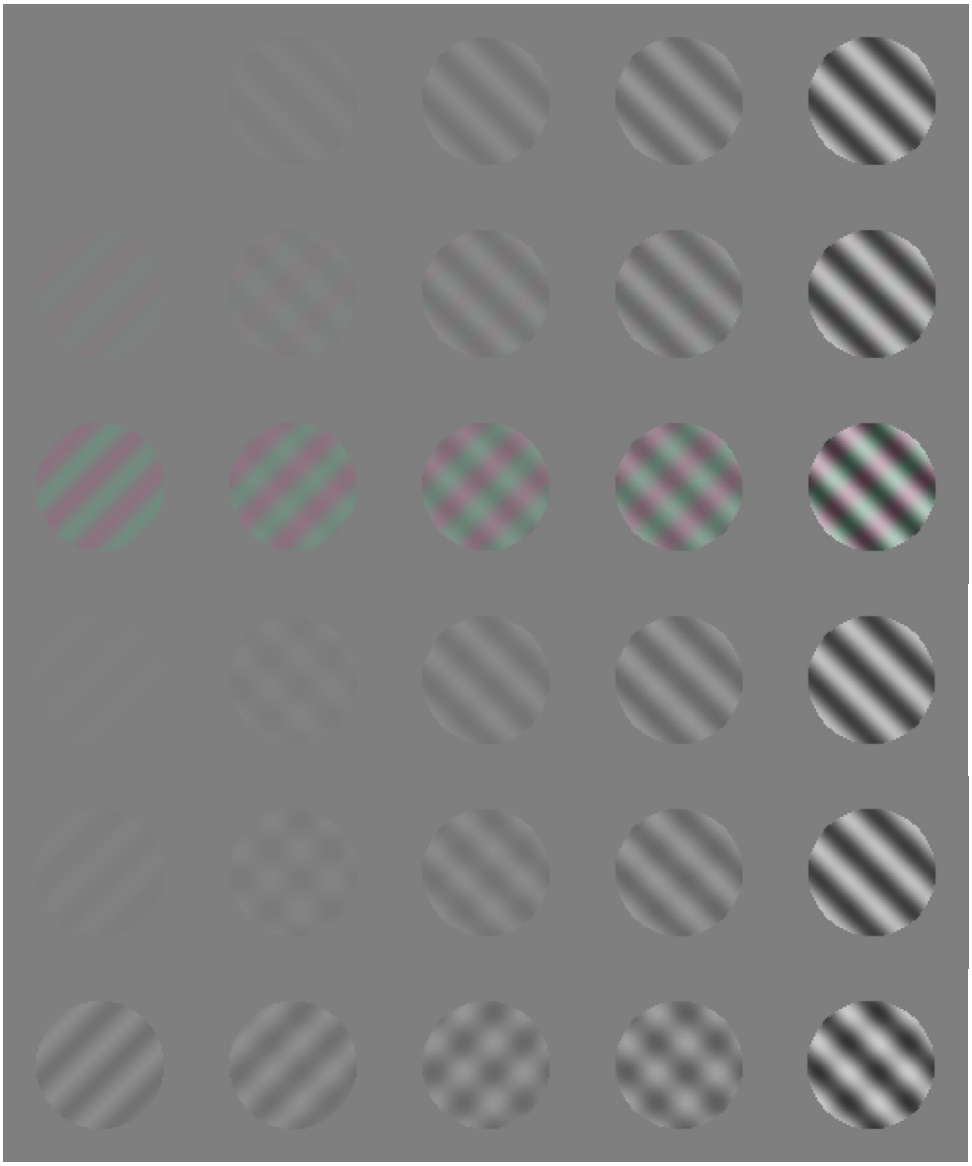


Figure 6.1: Representations of the 30 conditions used in the experiment. The order of the stimuli follows Table 6.2. Participants were asked to match the right oblique luminance component in all stimuli. The contrast adjustable matching stimulus was similar to the baseline stimuli (first row).

Data Analysis

I analysed each participant's data individually, as I expected from Chapter 5 to find individual differences regarding luminance and colour interactions. The data for each of the four sessions were grouped together and treated as a whole. This was done for each participant. The data obtained without masks is considered the baseline (I will refer to it as such in the rest of the text). Consequently, because I am interested in the modulatory effects of masks, I want to compare mask conditions with the baseline (no mask condition).

The baseline is different depending on the target contrast, so there are 5 baseline matched contrast levels, because of the 5 target contrast levels used. If the contrast of a

match is lower than baseline, I will describe it as suppressed; conversely, if it is greater than the target, I will describe it as being enhanced by the orthogonal mask.

From this data, I extracted the mean difference from the baseline for each mask condition, and then tested if this difference was significantly different from the baseline. For each target level, I compared the baseline level against the mask conditions using a t-test. As there are 5 mask levels per target, it is necessary to compensate for multiple testing (Bonferroni correction was used, shifting the significance level to $p<.01$ for conditions to be declared significantly different from the baseline). This analysis was performed for each participant and each condition.

Due to the number of participants and conditions, it is necessary to use a compact display, in order to summarize the data. This condensation should be performed while still trying not to lose any information in the process. In order to do that, I have displayed the results of each condition using a colour mapping of the difference from the baseline. I detail the procedure below.

I use a partial example for one participant to detail the visualization procedure of the dataset. Two conditions (the high colour and high luminance masks) were used as an example as well as the baseline for participant M4, but this procedure extends to all other mask conditions and participants. The raw data, of this subset example, can be seen in Figure 6.2, presented as a function of target contrast.

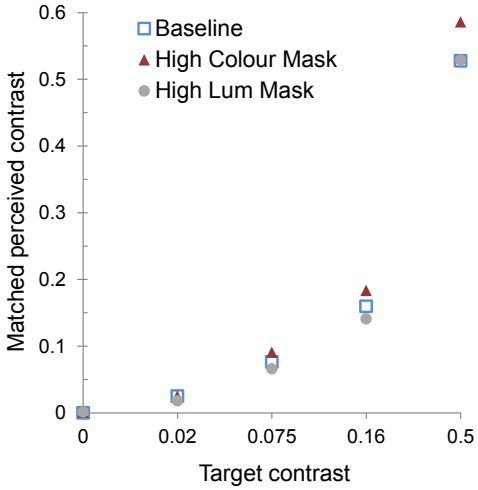


Figure 6.2: Raw data from participant M4 representing the mean perceived contrast for each target contrast for the baseline, high colour, and high luminance masks. The other conditions are omitted for clarity. The error bars (standard error of the mean), are smaller than the symbol used for data points.

Subsequently, the matched perceived contrast for each mask condition is subtracted from the baseline. This difference from the baseline can take negative (suppressive) or positive (enhancing) values (see Figure 6.3, left panel). The difference from the baseline is turned into a colour-map (see Figure 6.3 middle panel). The colour-map shows enhancing effects in orange and suppressing effects in blue. Each condition that was found significant

(Bonferroni adjusted t-test) is shown with a star. The data (matched values per contrast, test-baseline) for all mask conditions of one participant consists of a square (for participant M4, see Figure 6.3 right panel) that follows the same order used in Table 6.2 and Figure 6.1. The full set is presented in the following Results Section.

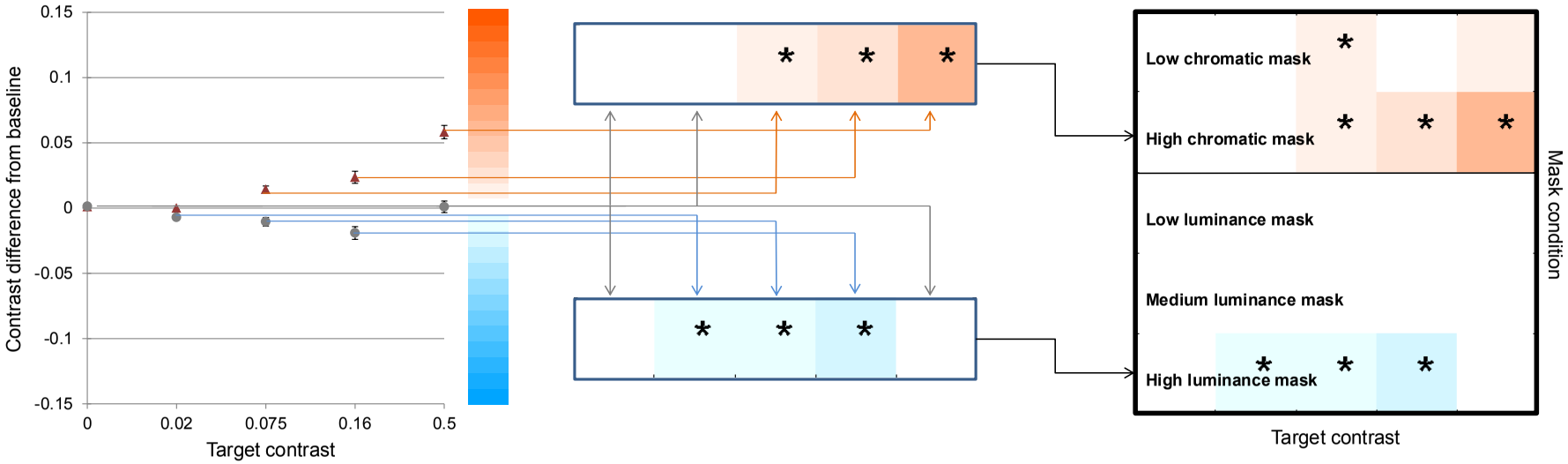


Figure 6.3: Data visualization. The data from Figure 6.2(participant M4), has been transformed into a difference from the baseline for high colour, and high luminance masks (left panel). An increase of perceived contrast is represented in orange shades and a decrease is represented in blue shades. The magnitude of the effect is reflected by the intensity of the shading. * denotes a significant effect using a t-test. The final square on the right side presents all the mask conditions tested.

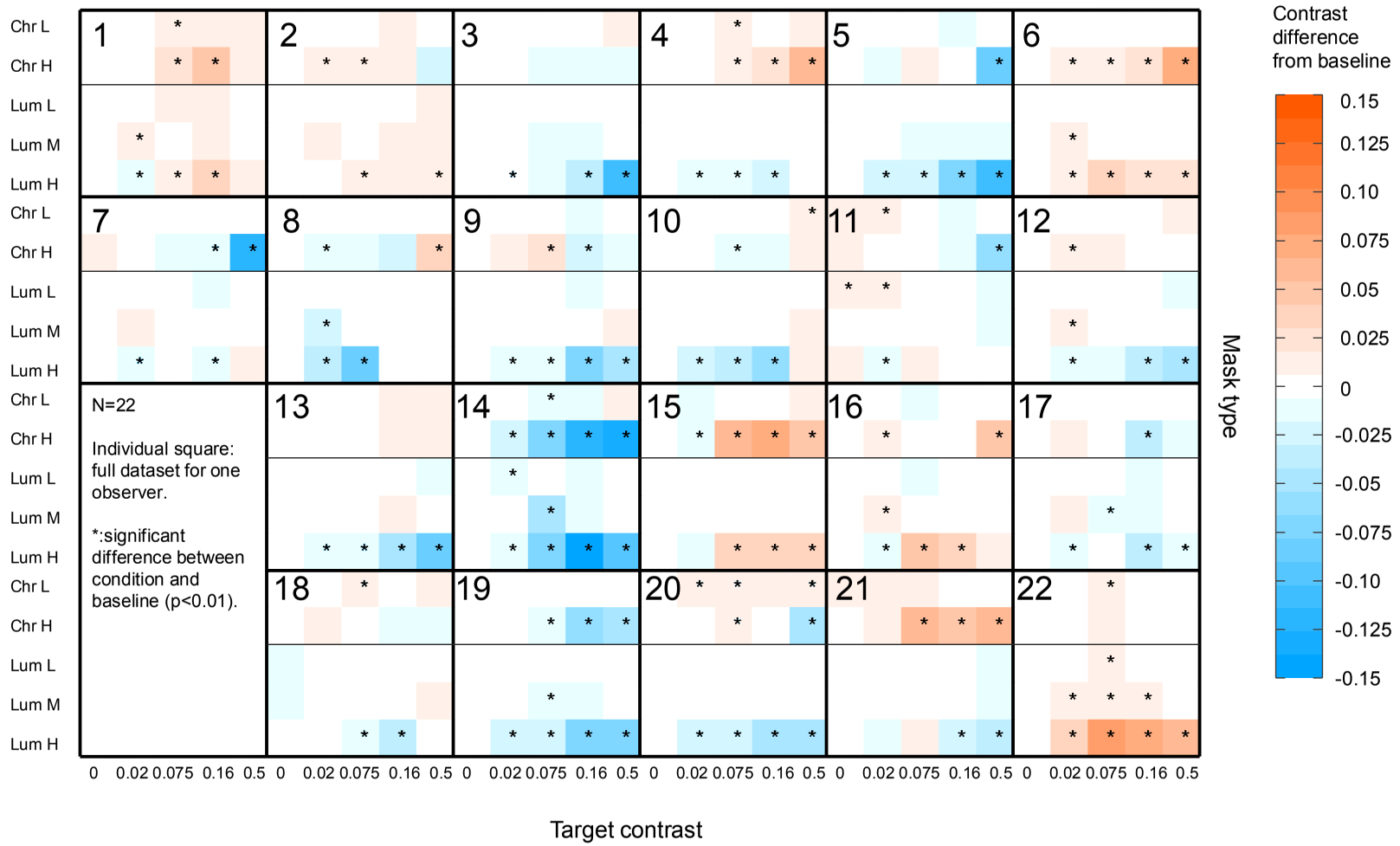


Figure 6.4: Observed difference between mask conditions and the baseline (no mask) for each participant, see Analysis section for details. Chr L: chromatic low contrast (0.01); Chr H: chromatic high contrast (0.03); Lum L: luminance low (0.01); Lum M: Luminance medium (0.02); Lum H: luminance high (0.10)

6.2.3 Results

All the results for every participant were processed according to the method outlined in the Analysis Section 6.2.2 and are presented in Figure 6.4.

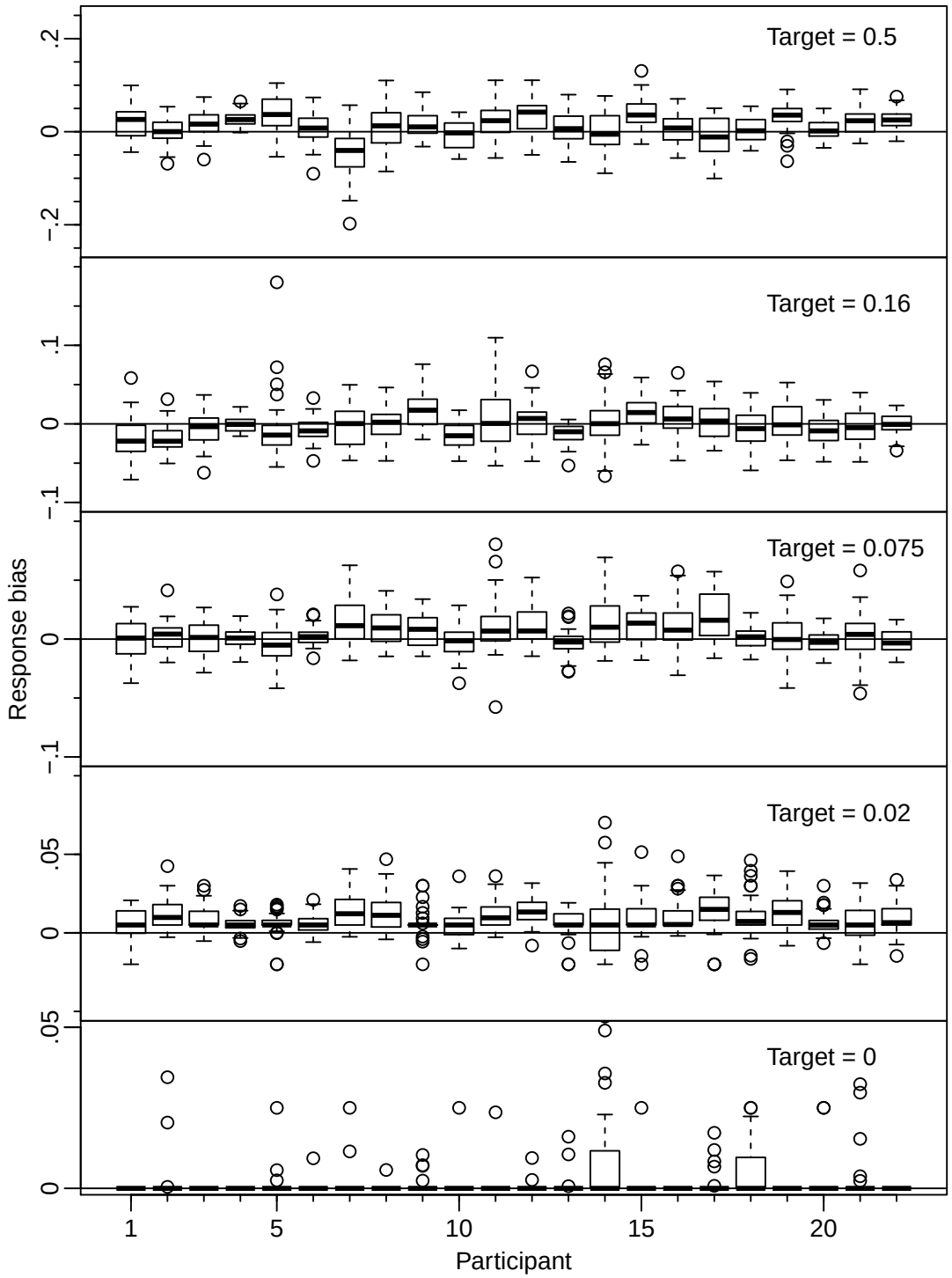


Figure 6.5: Observed difference between the baseline (no mask) and the expected target match for each participant. The difference should be around zero as there is no modulatory mask: bias would therefore be due to "response" bias. Bias is measured as match-target in (luminance) contrast values. Different y scalings are used for each target condition.

Baseline bias

I will firstly briefly discuss the baseline. In this chapter, the baseline condition refers to the mean target match when there is no mask present, i.e. when both colour contrast and luminance contrast are null in the orthogonal direction to the target. It is possible that some bias is present for this condition; some participants might tend to report more contrast (or less) consistently (see Figure 6.5). This is very much dependent upon the contrast of the luminance target. The biases, if present, tend to be at higher values of contrast. It is important to note that participants at 0% target (and in this Figure no mask) respond predominantly at zero (with few outlier trials, circles in Figure 6.5, probably due to pressing errors), showing that the participants were performing the task properly. Furthermore, at higher contrast values participants did match the target contrast with good accuracy and biases if present tended to be small. In Figure 6.5, the correct answer is represented by a horizontal line at 0 divergence from the target.

Here, I am generally interested in the divergence (Figure 6.4) of a mask condition from the baseline (seen in Figure 6.5). Another control condition was incorporated in the design of the conditions, the 0 contrast target. In this condition, it is expected that the baseline and the other masks are at 0 contrast. The first column in each participant's data summary should therefore be completely blank. This is generally the case for all participants with the exception of participant M11 in the low mask contrast (1%) condition. This could be interpreted as a misappropriation of perceived contrast. I therefore conclude that participants were performing the task properly and were responding to the luminance target in the correct direction, with possibly the exception of M11 who perhaps should be excluded from further interpretation for this low mask condition.[3]

Figure 6.5 & 6.4 (0% target), show that our participant could perform and understood the matching task properly. I will now have a preliminary look into colour and luminance effects independently, looking at mask and target tuning of the effects.

Colour modulation

Summary The main objective of this experiment was to look at the effect of colour mask modulations on perceived luminance contrast. To my knowledge, this data (Figure 6.4) is the first result showing colour contrast modulating luminance contrast perception.

The first noticeable property from the data is the range of behaviour displayed by the participants. This range spans from suppressive effects (e.g. participant M14, in Figure 6.4), to enhancing effects (e.g. participant M6), with some participants showing no (significant) effects (e.g. participants M3 or M13), and mixed/complex response patterns (e.g. participant M9).

Mask effect One pattern that should emerge from these interactions, as predicted by previous literature (e.g. Kingdom, 2003; Kingdom et al., 2005) and Chapter 5, is the

[3]Note however that participant M11 did not show strong biases in his baseline data, Figure 6.5.

expectation that there should be greater effect (either suppressive or enhancing) at higher mask contrast levels. Here the pattern of results in the data is not clear-cut regarding this prediction. However, a group of participants clearly showed this pattern of behaviour, participants M4, M6, M9, M14, M15, M19, and M21 (Figure 6.4, p. 135).

The variety of effect directions met the expectations from Chapter 5, as we are observing both types of effects, i.e. both suppressive and enhancing, as well as mixed ones. The implications of these results will be described in more detail in the Discussion section.

Target effect The second type of tuning expected should be luminance tuning. This is linked to the target contrast. Masking experiments show this type of behaviour (see for e.g. Meese & Holmes, 2007). More importantly, Kingdom et al. (2005) demonstrated this type of effects for the colour shading effect; the effect magnitude peaked when the luminance contrast was between 10 & 20% contrast, subsequently it dropped. I therefore expect the difference-from-baseline-effect to increase until around the 16% target and decrease subsequently.

The tuning is difficult to track down for two reasons. Firstly, there is a magnitude effect on the targets matched (here I am considering the baseline case, e.g. see Figure 6.5), meaning that there tends to be more variation at higher target levels. This effect is linked to Weber's law. Higher contrasts are less discriminable from nearby high contrasts than lower ones, consequently a small mean shift of the responses would be more difficult to detect.

Such target tuning (Figure 6.4) seems to be present for a small set of participants (M1, M9, M15, M17 and M19). For them the highest effect (negative or positive) is at the 16% contrast target. Another subset of 7 participants showed an increase of the effect (difference from baseline) as the target increases in contrast (M4, M5, M6, M7, M11, M14, and M21).

It can be difficult to disambiguate the source of this tuning. From the Baseline Bias Section above, we know that biases (variations from true contrast) tended to be higher at higher target contrast. This increased variation at higher contrast might hide small effects. Moreover, if one considers the processes responsible for inter-modulation between colour and chromatic channels, e.g. in models like the one by Meese and Holmes (2007), more complex patterns can emerge.

Luminance modulation

Summary There were several reasons to include the luminance masks. The main one was to compare the differences with colour masking. Additionally, the strength of the first two luminance masks were made low (1 & 2 %) in order to mimic the effects of small luminance artifacts that could have occurred in the colour masks if the stimuli were not equiluminant. If the colour stimulus contains luminance artifacts and these artifacts are responsible for the effects observed, then it should be possible to match the effects with low luminance mask conditions.

As stated in the Introduction section (2.4.4 & 3.3.3), usually luminance orthogonal masks are expected to suppress the target. To my surprise, the results show not only participants with suppressing but also enhancing interactions (identified in orange in Figure 6.4). However, Meese and Hess (2004) did report cases of facilitation in their contrast matching paradigm only in the binocular case (similarly to my protocol).

A total of 8 participants (M1, M2, M6, M11, M12, M15, M16 and M22) showed at least one significant enhancement using luminance masks, compared to 18 participants showing some suppressing effects.[4] A minority of participants are showing both effects (M1, M11[5], M12 and M16). Two preliminary results are of particular interest, one regarding the mask contrast and one regarding the target contrast. Most of these mixed effects appear only for one target condition at the target contrast value. Participants have either suppressive or enhancing effects at medium to high target values.

Mask effect Firstly, regarding mask contrast, most of the effects are found at highest mask contrast (lower condition rows in Figure 6.4). This might not be surprising on its own but it is interesting to see that 1% (lum low) and 2% (lum medium) contrast masks do not generate much modulation of the orthogonal suprathreshold contrast. In most of the cases in Figure 6.4 the panels are white, i.e. showing no differences from baseline. This has some implications for luminance artifacts, as I will describe later on in the Discussion section. The shorthand argument is that if luminance artifacts were responsible for colour masking effects, these luminance artifacts would need to be at least >2% to generate a significant perceptual shift (and this shift should follow the same direction).

Target effect Secondly, regarding target contrast, as stated earlier, I would expect some type of luminance tuning to occur. This is linked to the relationship between target and mask contrast. Work on luminance masking at threshold by Meese and Holmes (2007) showed suppressing but also enhancing effects in some conditions with target at threshold and using a lower contrast range. However, using contrast matching Meese and Hess (2004) did find both enhancing and suppressing effects (binocular condition only) and variability between their two observers regarding to when these happened.

The magnitude effect, whereby higher contrast targets have *more room* for variation, must be taken into consideration. It was a concern that effects might be masked by the variation of contrast at high values; however, the effects were large enough and the participants precise enough in their adjustment that the actual effects were not hidden. For example, at high target values strong effects are visible for participants M3, M14, and M22 (Figure 6.4).

There would be valid reasons to present the data as normalized by the target value, however I do not want to discard effects that might be due to magnitude directly. The

[4] All participants show signs of suppressing luminance with the exception of participants M2, M6, M15 & M22.

[5] See baseline, Section 6.2.3.

normalization does not take the discrimination differences at different contrast values into account .

The data (in Figure 6.4) hints at increasing effects with increasing target contrasts, with some participants (M3, M5, M12, M13, M19, & M20) having the highest effect at high target contrast, which would facilitate the fitting of the results. This is not the case for the whole dataset. Some participants (M9, M10, M14, & M18) show the kind of tuning described before for chromatic effects with increasing effects peaking at 16% contrast then dropping below the values for low contrast; another subset of participants (M8, M16, & M22) shows the same behaviour but with a peak at 7.5% target contrast. This tuning might be due to separability of contrast between mask and target (see Discussion).

Summary: Luminance and colour modulation

The pattern of behaviour described for luminance and colour masks requires more advanced analysis. I have discussed that both suppressing and enhancing effects where found for both luminance and colour masks. The strongest effects were found at the highest mask contrasts for both types of mask.

Consequently, I will now try to find a relationship between the luminance and colour results. This can answer several interesting questions such as the presence of a common orthogonal mechanism of spatial processing or possibly the use of common heuristics. Alternatively it could make more sense to parse the visual scene using different rules for luminance and colour as has been suggested, for example, by Kingdom (2003) (see Psychophysics Chapter, Section 3.6 for more details on this subject). Another alternative possibility would be that the luminance and colour results are linked because the chromatic stimuli have luminance components within them (i.e. artifacts); I do not believe that this is the case, for several reasons, which I will describe in the Discussion section.

The next section investigates the commonality of results between luminance and colour masks in a systematic way for my whole population tested. If a common pattern is found this would indicate that one of the hypothesis described above might explain the observed pattern of results.

Similarity between chromatic and achromatic modulatory effects

Introduction In this analysis, I am looking into how colour and luminance show similar or dissimilar effects at the population level. The fact that we are dealing with a group of patterns has direct implications for the interpretation of the results. One should be cautious to interpret results from a small sample of participants and infer that the whole population is behaving with the same pattern of behaviour. This is a problem that is common and not frequently discussed in the literature because it is difficult to address with psychophysical experiments requiring many of testing hours.

The number of participants tested in this experiment is larger than in many previous experiments but still constitutes a small sample of the population. The homogeneity of

the patterns of responses within that sample is of interest. In the rest of this section, I will detail several hypotheses on the distribution of these effects (suppressive/enhancing) across the population. Here, I present a list of possible patterns that one might expect from the simplest to the more complex.

Prior to the experiment, some hypotheses were formulated. Firstly, one could expect to see only suppressive effects from orthogonal masks. This is clearly not the case as it can be seen from the results presented in the previous section and this comes from the fact that usually orthogonal luminance effect are supposed to be suppressive. This finding is mostly from detection tasks at threshold level; however, see Psychophysics and Physiology Chapter 2 & 3 for a more nuanced review showing enhancing effects in specific conditions.

Another hypothesis suggested by the results from the previous experimental chapter on colour modulation would be that colour masks show suppressing, enhancing or no effects, with an always suppressive luminance interaction.

I now detail hypotheses regarding the variety of observed effects. One could expect to find correlated luminance and colour effects within individuals, suggesting that the suppressive effects are triggered by a similar mechanism, or at a higher level suggesting that the participant is using a similar *strategy* or heuristic to solve the task.

Another possibility would be that that the effects are anti-correlated, this would fit with the scene analysis hypothesis (Kingdom, 2003) where luminance and colour play different roles in segmentation processing. Here, I am still discussing effects at the population level, so participants might have an increase/suppressive pair (i.e. showing both types of effects) that could be either chromatic/luminance or the opposite as a general tendency.

Finally, the possibly less interesting hypothesis is that there is no general tendency within the population, which could be interpreted as a lack of functional role for these interactions (at least in the suprathreshold contrast perception domain).

As a summary, discarding the first two hypotheses, the results at the population level, fall into three categories. The results are correlated, with both luminance and colour having the same effects for each observer, anti-correlated, or uncorrelated. The theoretical implications of those cases are described in more details in the Discussion section of this chapter.

Analysis In order to distinguish between the possibilities described in the previous section, I looked at the relationship between the effects of the highest luminance mask and highest colour mask contrast. This was chosen because this is usually where the highest effects were found as can be seen in Figure 6.4. Looking at Figure 6.4 it is difficult to see if the direction of the effects between luminance and colour tend to be similar within the population.

The relative distribution of difference from baseline for both colour and luminance (high contrast) masks are presented in different forms in Figures 6.6. Figure 6.6 presents the averaged values. Each average is calculated from 40 trials (sampled over 4 sessions) and as such it does not take into account the variability of the data.

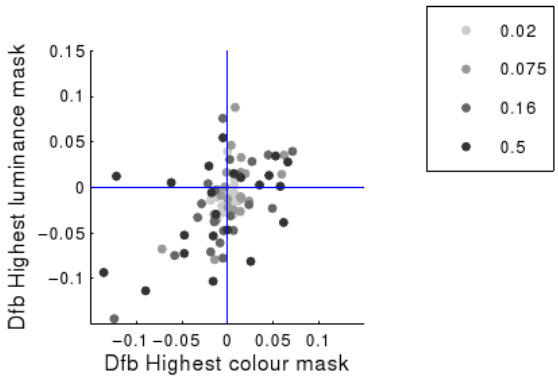


Figure 6.6: Luminance mask effect versus colour mask difference from baseline (Dfb) effects (% luminance contrast). For all observers the mean difference from baseline effects for the highest luminance and colour masks are presented against each other. Each target contrast is presented by a different filled circle shade.

These summary data clearly show that there are no purely suppressive effects; all the results would be distributed in the lower left quadrant of Figure 6.6. As stated earlier it was obvious from Figure 6.4 that this is not the case. The second hypothesis (suppressive luminance effect and variable colour effects between participants) would have led to most of the variation being distributed in the colour dimension, effectively showing all of the data in the lower two quadrants. As is obvious from Figure 6.6, this is not the case either.

Interestingly, as stated earlier, the luminance masks can have both enhancing and suppressing effects within the population. Here we are left with three hypotheses. In order to look into the relationship, I analysed the density functions of these two effects.

Density For further data exploration, the data were transformed into a density map (Figure 6.7 c), which is a proxy for the joint distribution of both colour and luminance mask effects. The density is calculated using binning. Essentially, I created a MatLab script to count the number of data points present in each bin (see Figure 6.7-d). The number of bins is selected by the author and can be adjusted. By dividing the bin content by the total number of points, I obtain the probability surface of the combined effects, or joint probability (Figure 6.7-a& -d, side and top view). Figure 6.7-c uses a heat map to represent density and Figure 6.7-d presents a side view of the density map. These plots show the distributions of difference from baseline effects at the highest luminance and highest mask effects for all target values. It is evident that the distribution centres around zero; this can also be seen on the marginal distribution for luminance and colour masks (Figure 6.7-a, -d). The shape of the distribution is important as it will tell us which hypothesis is more likely regarding the relationship between luminance and colour processing. The densities are calculated from a set of 22 participants for 4 target contrasts (88 points in total). Each point is the result of 40 trials in the highest luminance mask (Lum H in Figure 6.4) and 40 trials in the highest chromatic mask (Col H in Figure 6.4).

The marginal distributions of both individual effects can be seen in Figure 6.7-a&d.

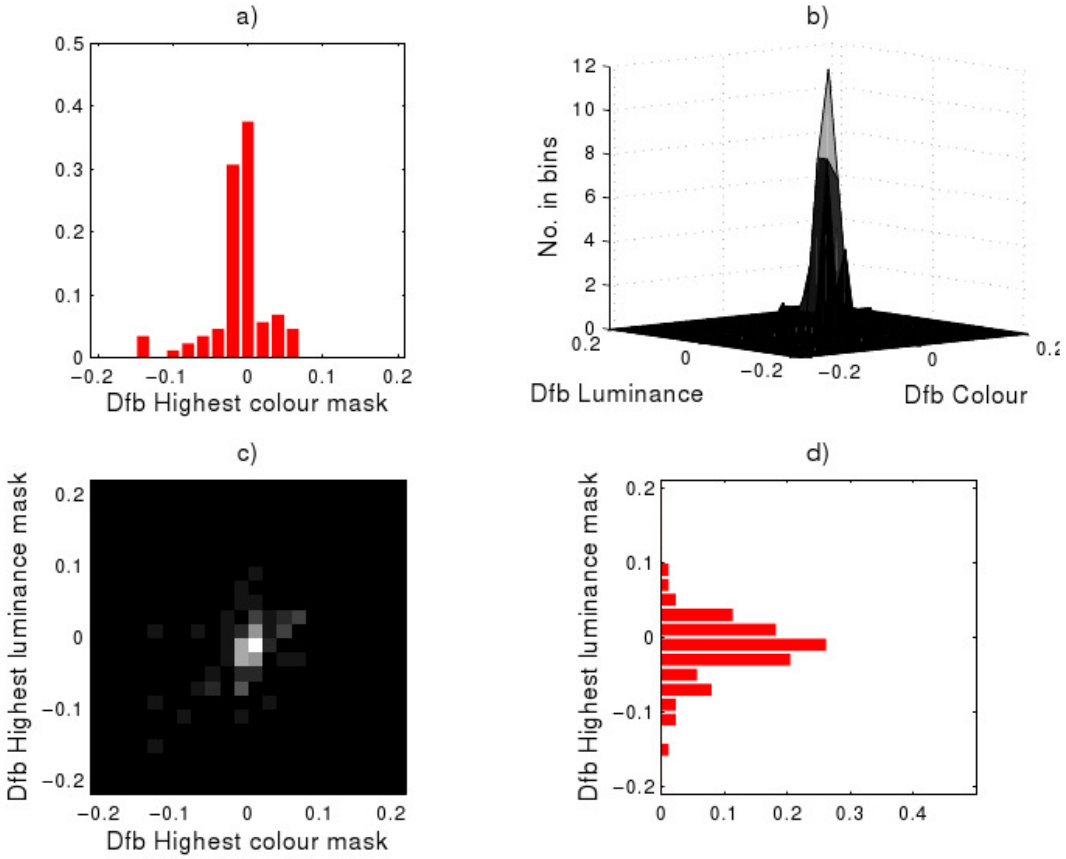


Figure 6.7: Density of luminance mask effect versus colour mask difference from baseline (Dfb) effects (% luminance contrast) a) (Marginal) Probability density of colour masks effects for all targets contrast values and all participants, independent of luminance effects. b) Side view of the shared density of difference from baseline effects. c) Top view of the joint probability of difference from baseline, this is essentially a density map of a). d) (Marginal) Probability density of luminance mask effects for all target contrast values and all participants, independent of colour effects.

They represent the individual distributions of effects for luminance and colour, at high mask contrasts. From these two marginal distributions, it can be seen that the luminance has a stronger effect (at the level of the population) than the chromatic mask. Its distribution is wider. Both distributions seem centred around zero or have a slight bias toward suppressive effects (negative values on Dfb graphs). Note that, I have used lower mask values for luminance and colour; the distributions would have been tighter and centred around zero.

Correlation and mutual information The next step of analysis is to check how much one effect is linked to the other one. It is possible to find the mutual information using the data from joint and marginal probabilities. The **mutual information** gives a metric of mutual independence. For two discrete variables X and Y, the mutual information is calculated (see Hansen & Gegenfurtner, 2009, for an example on natural image statistics)

using the following equation:

$$I(X;Y) = \sum_{y \in Y} \sum_{x \in X} p(x,y) \log_2 \left(\frac{p(x,y)}{p_1(x) p_2(y)} \right) \tag{6.1}$$

where p(x,y) is the joint distribution (see Figure 6.7-c) and the marginal distributions (p(x) and p(y)) are the differences from baseline for luminance masks and colour masks, respectively (Figure 6.7-a& -d). The maximum value of mutual information (MI) is dependent upon the number of bins selected to calculate the joint density functions and its histogram counterparts for the marginal distributions. The maximum value that can be obtained is $log_2 N$ where N is the number of bins.

Alternatively, one can use the **correlation** between the two variables and use the correlation matrix to fit a bivariate Gaussian. These metrics can be obtained quickly using the information shown in Figure 6.7. The overall correlation is .53 and the mutual information is 0.795. The maximum value of mutual information that can be obtained, linked to the number of bins, is $\log_2(22) = 4.46$; values close to zero would indicate independence. These two values indicate a correlation between the two effects.

However, there is a distortion in the dataset that needs to be taken into consideration. Looking at Figure 6.4-a& -b, it can be seen that the effect strength will spread more or less around zero depending on the target contrast.

As an initial data exploration, all target contrast values were plotted together to check for a global pattern in the results. However as it can be seen in Figure 6.6 -a& -b and as can be expected at lower target values the distortion is less. The lighter points indicating low target contrast tend to cluster more around the centre than the darker points (6.6 -a& -b). This intuitively makes sense, as there is more possibility for modulation as magnitudes increase; it is the *magnitude effect* mentioned in earlier sections. This type of effect can be expected when the type of modulation is not additive but multiplicative. However, the tuning of the modulation with the target is a more complex issue and will be discussed later on.

Analysis by target contrast In order to avoid the issue of different strength effects for each target contrast, I *sliced* the dataset used previously by target contrast values. Consequently, I obtained the joint and marginal distribution of colour and luminance effects (or differences from baseline) for different contrast values. The *sliced* dataset can be seen in Figure 6.8, with each row representing a different target contrast. The first column shows the joint distribution at each target contrast, in grey scale. The second column shows the associated bivariate Gaussian fit (also implemented using Matlab). In graphs in the first column the origin is shown by a blue cross and a square is displayed with a radius of the target level. If a participant perceives the target with no bias in the baseline condition (i.e. veridical contrast perception), and then shows an effect that fits on the border of the blue square, it means that the effect of the mask on the target is of the magnitude of the target; whether it is suppressing or enhancing.

It is important to notice that the extent of the distribution is somewhat bound by the target level. For example in the case of a 16% target, I would not expect to have suppression of more than 16%.[6]

In the second column of Figure 6.8, the parameter estimates are directly obtained using the correlation factor p and the correlation matrix of the data. The third and fourth columns display the marginal distributions; two blue bars are displayed on the marginal distributions representing $\pm$ target values (similarly to the square in the first and second column).

Note that there are only 22 data points that are taken into consideration for this density calculation and subsequent correlation. For each target value, I calculated the correlation and mutual information.

The results of these analyses are presented in Table 6.3. It appears that at higher target contrast values the results are highly correlated. As was expected the lowest target value (2% contrast) is pulling the correlation down. The highest correlation is at 16%; that can be seen easily in Figure 6.8-c by the elongated shape of the bivariate and lower left data point in the first column. This data point is close to the blue square, meaning the participant is reporting an almost completely suppressed perception of contrast with those masks, although this appears to be caused by one point only (consequently one participant), who displays strong suppressive effects for both types of masks. The mutual information indicates that the shared information *grows* from lowest to highest contrast values.

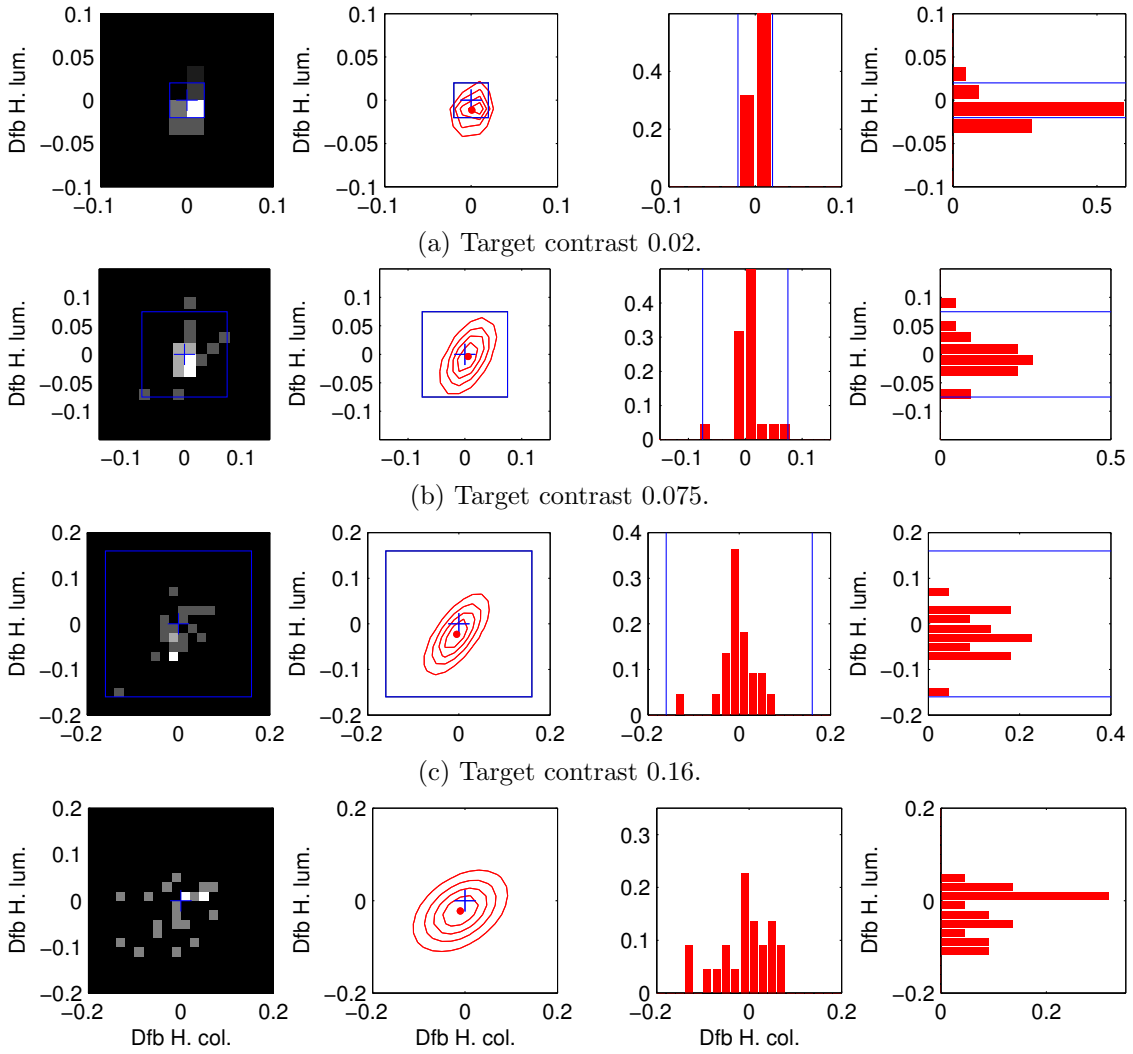
A high correlation (or shared information) would suggest that participants tend to have similar behaviour between chromatic and luminance masks. However, due to the small number of data points used to calculate these metrics the results are prone to be distorted by outliers. This is especially the case for target values of 0.16 (see Figure 6.8-c, one participant (M14) is clearly driving the correlation).

The correlation found in this first analysis would suggest the existence of common mechanisms driving these effects at the level of the population (and not for all participants). However, as noted previously, this analysis only takes into account the mean of the responses for each participant. It is thus ignoring the variability of the modulatory effect. This will be taken into consideration in the follow up analysis presented in the next section in an attempt to remedy this shortcoming.

Extended analysis: Colour and luminance effects

Introduction One possible critique of the analysis conducted so far is that each of the correlation (or independence) metrics is based upon only 22 data points, one from each of our 22 subjects. As I am using the mean, I am also brushing away the variation of the dataset; this should be taken into consideration. Formally, for each point presented in

[6] It can however happen as we are here dealing with differences from baseline. For example, if an observer has a bias at the baseline (Figure 6.5), if this tendency is to respond more and if in the mask condition the stimuli are completely suppressed, then we would obtain a suppression >16%.



(a) Target contrast 0.02.

(b) Target contrast 0.075.

(c) Target contrast 0.16.

(d) Target contrast 0.5.

Figure 6.8: Colour and luminance masks analysis by target values. From top row to lower row luminance target values 0.02, 0.075, 0.16 and 0.5. First column: joint probability density. Second column: bivariate Gaussian fit. Third column: colour masks marginal probability. Fourth column: luminance mask marginal probability.

Table 6.3: Table of correlation metrics between high chromatic and luminance masks per target values.

	Target contrast			
Measure	0.02	0.075	0.16	0.5
Correlation	0.16	0.55	0.68	0.42
Mutual information	0.104	0.690	1.261	1.719

Figure 6.7 b-c, there are 40 independent measures taken in both luminance and chromatic dimensions of the graph. This variation should somehow be taken into consideration. However, data points are collected independently, in separate trials; it is therefore difficult to know how to pair the 40 by 40 trials for colour and luminance. One general rule for analysis, which I have tried to follow as much as possible in this thesis is to make as few assumptions as possible. In the following section, I present a method to use all the data points without making any combinatorial assumptions.

Analysis I found there were two possibilities to deal with this issue. One approach would be to take samples from the 40 points distributions, do this many times and see how the measures described above vary, i.e. *bootstrapping* the dataset. Another possibility is to display all the data points against each other making no assumptions about their relationship. Consequently each point previously displayed in Figure 6.7 becomes a 40 by 40 matrix of points of all possibilities. This method is not based on bootstrapping, per se, as it does not depend on repetitive sampling but it will give us an idea of the spread of the data, which as I said is lost with the use of means. Consequently, I do not expect an increase of correlations; this is a first test of the robustness of the correlations found above. The second method was used.

Results Figure 6.9 shows the results of this modified analysis; it is using the same layout as Figure 6.7. In Figure 6.7b-c the original data of 22X4, 88 points (participants by target levels, averaged per condition) is replaced by 22X4X40X40, 140800 data points (the resulting cloud of points is not shown). Regarding the number of points displayed, it is more telling to look at the density plots (Figure 6.9c & b). The correlation on the whole dataset is .24 and the mutual information is 0.178. In this case 52 bins were used so the mutual information is bound to a maximum value of $\log_2(52) = 5.70$. This also applies to the subsequent analysis performed on the *sliced* dataset. Similarly as in Figure 6.7 it is evident that there is a peak around the centre of the distribution. This might be driven by the lowest target contrast.

Consequently, I performed the same analysis as before by separating each target's contrast independently. For each target contrast, I calculated the mutual information and correlation. The results are presented in Table 6.4. Comparing Table 6.3 & 6.4, it can be seen, as expected, that the correlation is generally lesser in the extended dataset analysis. However, the relative pattern between the targets is similar. The third target level (0.16 contrast) has the highest correlation level (0.36) and the lowest target contrast (0.02) has the lowest correlation (0.05).

The *sliced* dataset is presented in Figure 6.10, using the same convention as Figure 6.8. Looking at the marginal distributions (column 3, 4) one can see the effect of using all the data points (in a 40X40 combinatory fashion). The distributions were effectively smoothed. The distribution seems to be long-tailed, allowing more extreme values to occur. Nevertheless, the bivariate Gaussian fit seems to capture the distribution well.

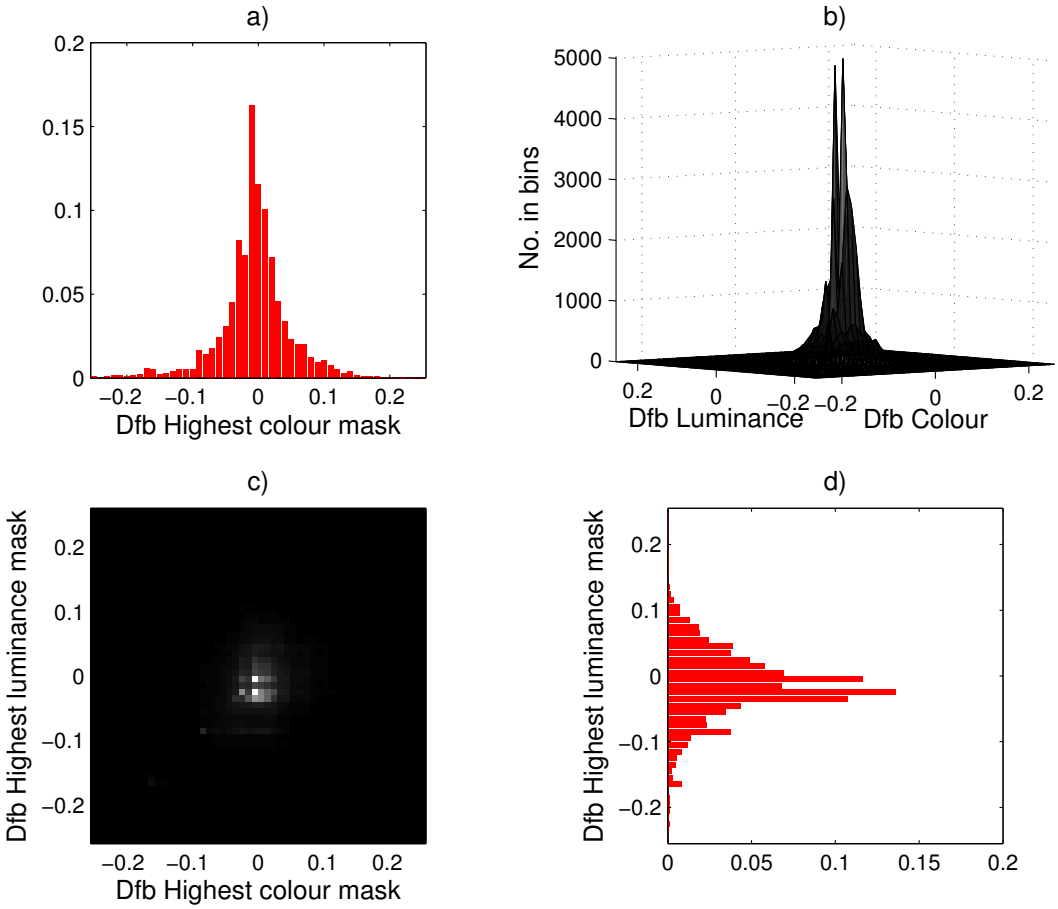


Figure 6.9: Luminance mask effect versus colour masks effects for all trials. a) (Marginal) Probability density of colour mask effects for all target contrast values and all participants, independent of luminance effects. b) Side view of the shared density of differences from baseline effects. c) Top view of the joint probability of the difference from baseline (see text). d) (Marginal) Probability density of luminance mask effects for all target contrast values and all participants, independent of colour effects.

Once again, the same participant (M14) is pushing the correlation up (as tested by performing the same analysis without this participant, data not shown). It is possible to spot that participant's data in the marginal distribution in target 7.5% & 16% (see Figure 6.10-b&-c, marginal distribution for luminance and colour high contrast masks in the third and fourth column). It corresponds to the spike in the data at the negative outer bound. It was considered whether this participants' data should be discarded from the dataset. However, looking at the data in Figure 6.4, it can be seen that this participants (M14) consistently perceived less contrast in the presence of masks. Some participants such as M3 and M7 only showed similar magnitude of effects at the highest target contrast and in only one mask type. Participant M14's data should not be discarded based on having the same effect in both luminance and colour masks because this also happens for others. For example, this is also the case for participant M19 but with less effect strength. The possible reason why participant M14 is having such extreme behaviour will be explored

Table 6.4: Table of correlation metrics between high chromatic and luminance masks per target value for the full dataset (see text)

	Targets			
Measure	0.02	0.075	0.16	0.5
Correlation	0.05	0.29	0.36	0.16
Mutual information	0.091	0.168	0.294	0.117

further in the Discussion section (especially regarding the use of heuristics). Note that with participant M14's data discarded the values of correlation are .04, .19, .21 and .12; which can be considered low to medium.

In this section, I have described that there was a correlation between the luminance and mask effects; however, this correlation became small when using the whole data points for these conditions (see Extended analysis). This was especially the case when participant M14 was taken out of the analysis. From the hypothesis formulated, the anti-correlated version has no basis. Depending on the analysis procedure used, there is either evidence for correlation or no-specific effect. In the case of a correlated mechanism, when one effect is low in one mask type we would expect it to be low in the other one. However, there are a lot of examples in the raw data when only one type of mask is really efficient (e.g. 3, 11, 13, 18) with the other mask type having no effect.

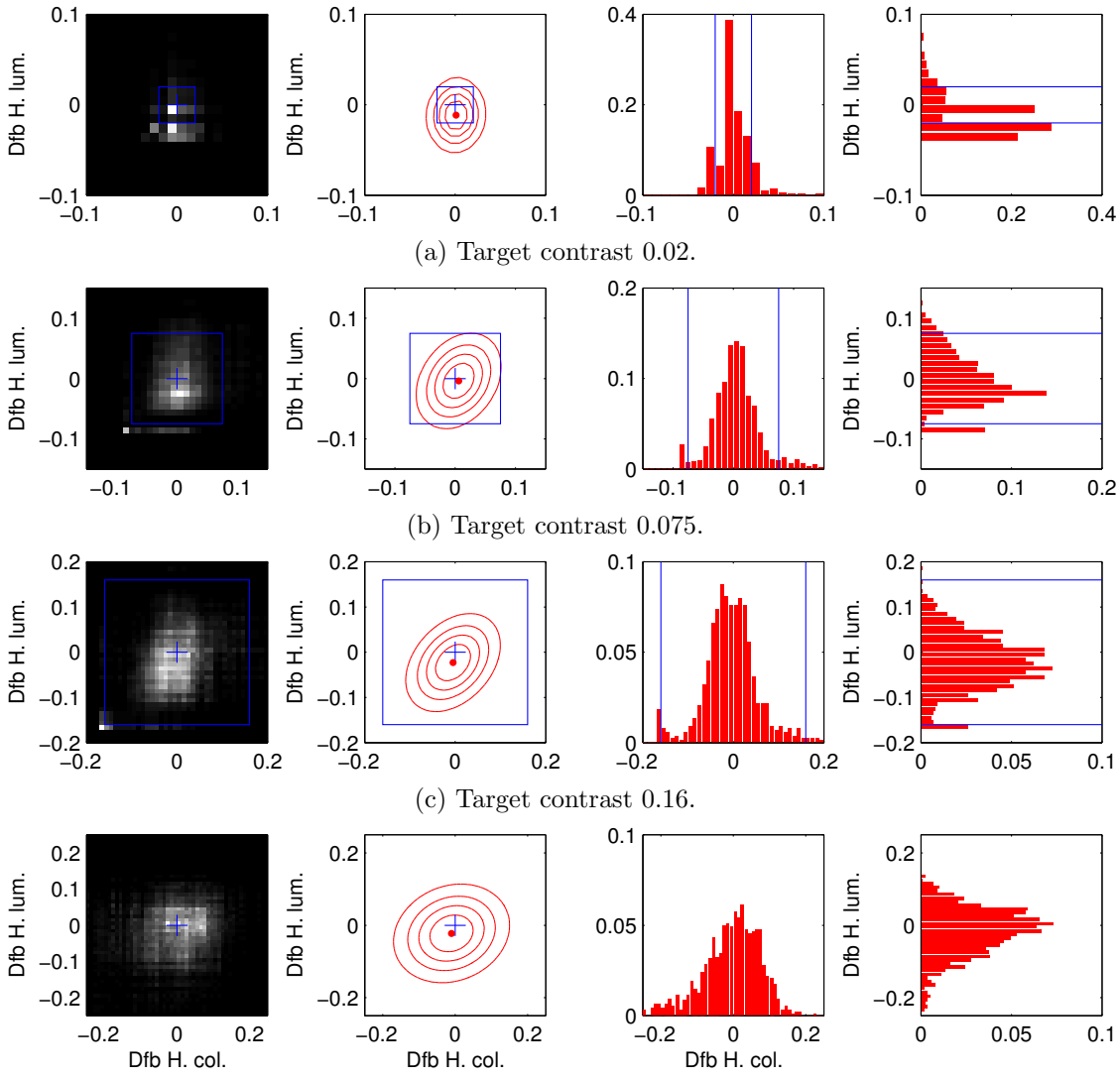
The evidence seems to favour a non-specific correlation at the population level. To be clear, that does not prevent individual participants from showing correlated or anti-correlated effects (or suppressive/enhancing effects only for one mask type), as was actually found in this dataset. This leads to three questions: the functional relevance of those interactions, how they come to be in place in the visual system, and finally, if these effects extend to other percepts. In particular, does it link to shape-from-shading and colour shading differences observed in the previous Chapter 5? These issues will be discussed in Section 6.3.

In the next section of the results, I will look into the potential shifts in the data that could be due to adaptation or training effects.

Consistency analysis

Here, I will be looking into adaptive effects that might be occurring within sessions and during the whole experiment. They are two main types of effects that could occur during the experiments.

The first type is due to adaptation during a session. As it is a free viewing experiment, one could expect some adaptation effects to occur for prolonged exposure. However, the eye movements between the target and the match might play against this. Secondly, the fact that it is a matching task and not a detection task might mitigate any adaptation effect, as it would affect both the target and the match stimulus equally. The second type



(a) Target contrast 0.02.

(b) Target contrast 0.075.

(c) Target contrast 0.16.

(d) Target contrast 0.5.

Figure 6.10: Colour and luminance mask analysis by target values with the extended dataset (see text). From top row to lower row luminance target values are 0.02, 0.075, 0.16, and 0.5. First column: joint probability density. Second column: bivariate Gaussian fit. Third column: colour mask marginal probability. Fourth colour: luminance mask marginal probability.

of adaptive effect would be learning effects across the sessions. This type of drift would show an upward or downward trend across the trial number.

In order to check for a drift, I looked into the difference from baseline effects for the whole population against trial number. This was performed for the high luminance masks case, as I expect stronger adaptation effects to be found with the stronger contrast presented. The results are presented in Figure 6.11.

Figure 6.11a shows the difference from baseline as a function of trial number. There is no apparent drift in the data from trial 1 to 40 in the population as a whole. I chose the highest values of target contrast to check for drifts as it should be more apparent at this level (regarding the magnitude effects discussed earlier). Secondly I chose the highest luminance masks as it is where strongest effects were found (see Figure 6.4). Essentially, if these effects were due to adaptation, then this could be seen in the trial-by-trial plot shown in 6.11a.

The distribution of the difference from baseline effects is actually the same marginal distribution displayed in Figure 6.10d.[7] The lack of an apparent general shift could suggest that no learning occured during the sessions.

For the second type of adaptive effects, I was looking into adaptation during each session. This kind of effect would look like a saw-tooth pattern in the trial data (Figure 6.11a) with the pattern re-setting at every new session. This was not evident in the data. However, in order to make a more thorough examination, I performed a half-split analysis on the data. The first half of every session for every participant were combined together and compared to the second half. The distribution of the difference from baseline as well as a box plot summary are presented in Figure 6.11b. Here, again, it is clear that no major shift seems to explain the distribution of results obtained in the dataset.

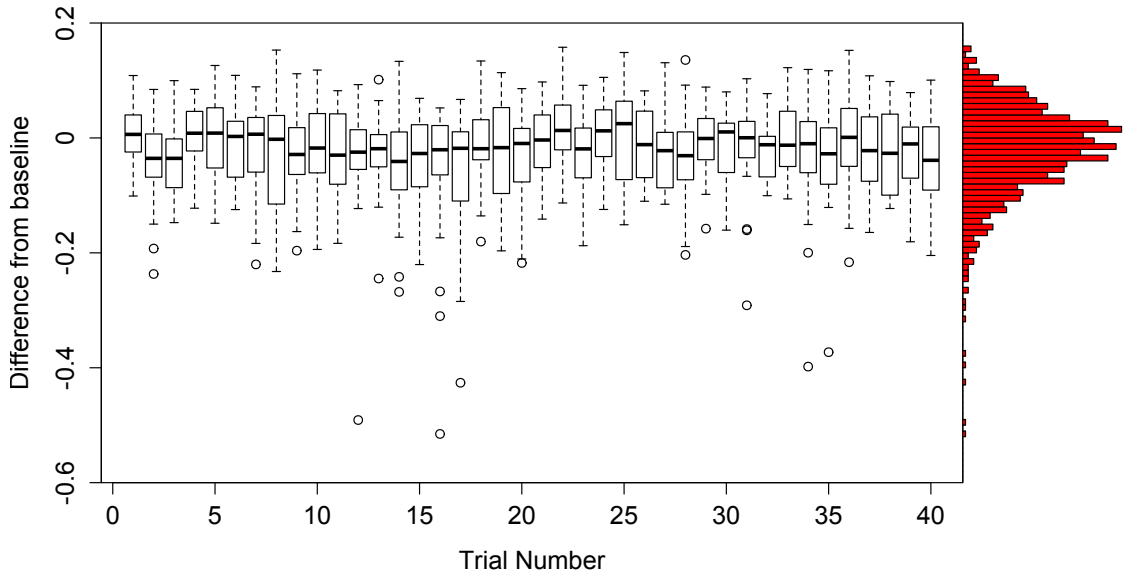
The same analysis was performed for the highest target contrast at the highest colour mask with similar results (see Figure 6.12). I am confident that these are indicators that no adaptive/dynamic effects (across session or across the whole experiment) can explain the results.

6.2.4 Comparison with colour shading results

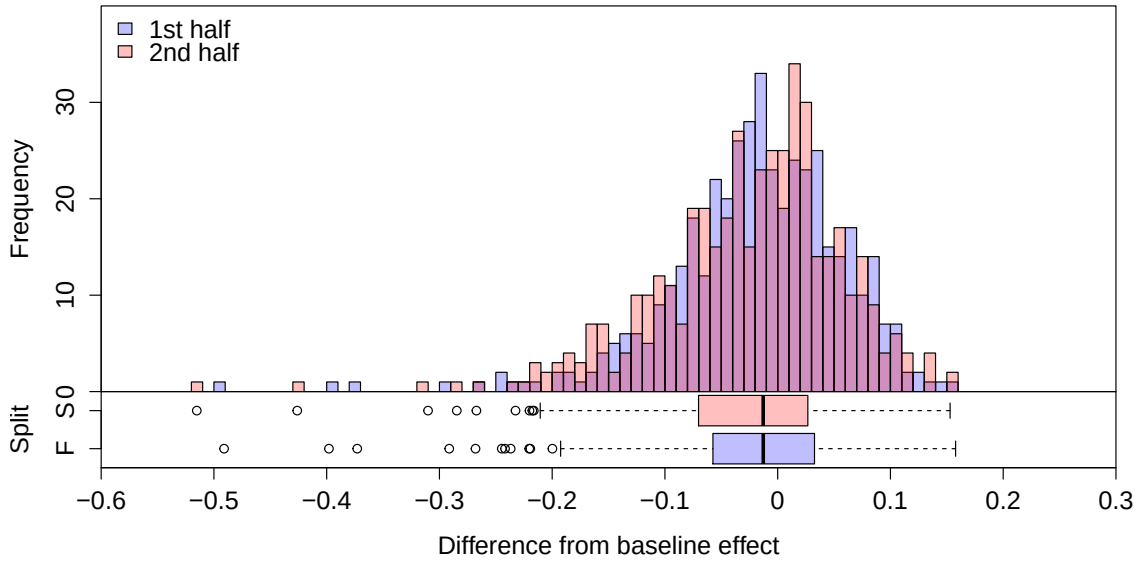
Introduction

One of the aims of the contrast matching experiment presented in this chapter was to find out whether the variations obtained in the colour shading experiment (Chapter 5) could be due to low-level effects in contrast perception. The methodology and stimuli used in the luminance matches varied from the first experiment but some conditions do overlap to some extent. The most important overlap is in the subgroup of participants that were tested in both experiments. This allows for a more in-depth analysis of their contrast and depth perception (as well as isoluminance perception, see Chapter 8 for this analysis).

[7] Differences in binning explain the marginal differences in the observed shape.

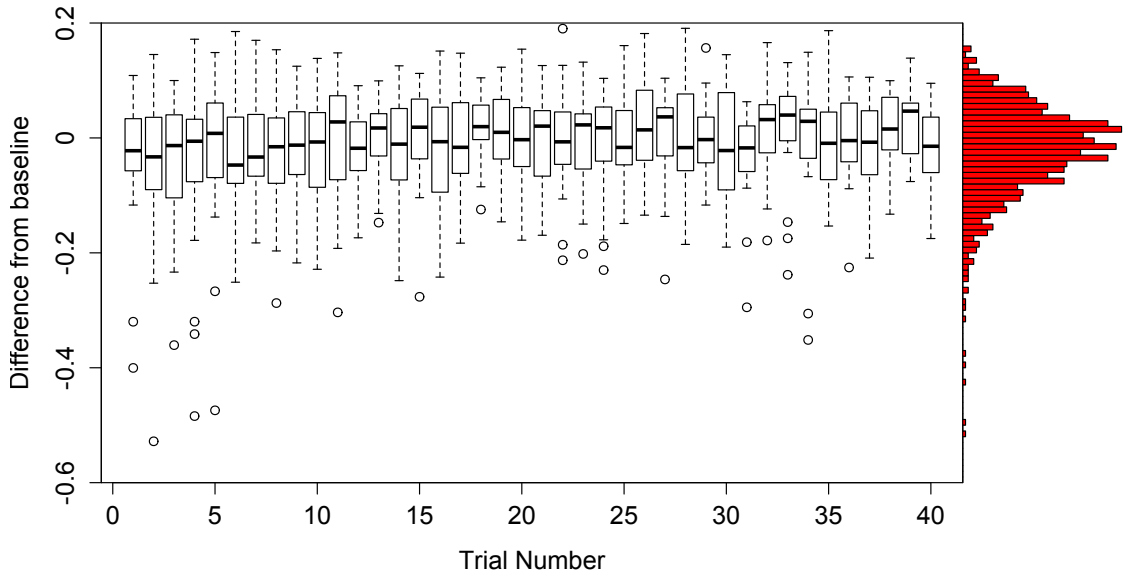


(a) Distribution of difference from baseline effects for the participant population by trial number.

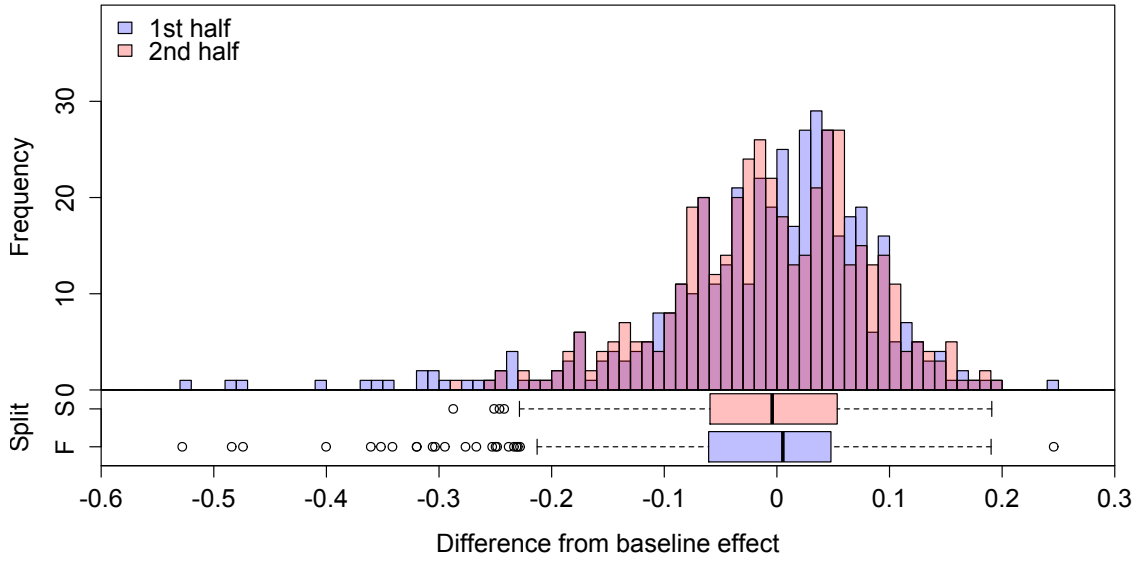


(b) Distribution of difference from baseline effects for the first half and second half of each session.

Figure 6.11: Drift analysis of all participants' data for the high luminance masks. For each participant data was taken as difference from baseline.



(a) Distribution of difference from baseline effects for the participant population by trial number.



(b) Distribution of difference from baseline effects for the first half and second half of each session.

Figure 6.12: Drift analysis of all participants' data for the high colour masks. For each participant data was taken as difference from baseline.

I will first list the major similarities and dissimilarities between the stimuli and protocol of the two experiments in order to frame the results more accurately in their context. Following that, I will detail some hypotheses that were suggested on how contrast perception might influence the shape-from-shading experiment. Finally, I will present a summary of the results from both experiments, for the participants who participated in both.

Protocol differences

Stimuli The stimuli used in both experiments were the same in terms of spatial frequency and spatial extent; additionally, the component orientations were the same. However, in the contrast matching experiment a Gaussian window was used at the edges instead of hard edge windowing.

The colour shading experiment involved both sinusoids and square-waves whereas the patterns used for the colour matching experiment were sinusoids only. Consequently, in the following, I compare only the results for the sinusoids (the same approach was taken for the isoluminance comparison in Chapter 8, Section 8.4.3).

Contrast values In terms of contrast used, both luminance and colour were different. However, the colour contrast spans a similar extent in both experiments and it can essentially be plotted out as such. In the colour shading condition 2, the luminance was fixed at 4%. In the matching experiments the closest target contrasts tested were at 2% and 7.5%. Due to this discrepancy and due to the delay in testing (>12 month) we have to remain cautious in the interpretation of the results.

Procedure Both tasks used in these experiments were matching tasks (*method of adjustment*) with continuous free-viewing. The time for each session was similar. All participants first performed the depth matching experiment and so the data are not properly counter-balanced. One could possibly argue that a learning effect appeared due to the performing of the task. This would however suggest that performing two hours of depth matching has a lasting effect on contrast perception after several months; I consider this unlikely.

Hypothesis

It was suggested in the introduction of this chapter that modulation of depth from shading by colour, or the colour-shading effect, could result from interactions in the processing of contrast before the encoding of depth. This is a purely bottom-up approach to the phenomenon. If such interactions were present, then I would expect increases in perceived depth to be due to reduced luminance contrast and the opposite for a decrease in depth. Note that this assumes that it is possible to probe contrast perception and shape perception independently. At a physiological level it is reasonable to expect population encoding for these features alone.

However, at the perceptual level it might be more difficult to separate both percepts. For example, in Bloj et al. (1999) showed that the perception of the shape of the object influences the perception of other basic properties (colour in this case). This is essentially a case of top-down influence of the scene interpretation on the reflectance estimation. This issue will be discussed in the light of the results (presented below) in the Discussion section of this chapter.

Results comparison

The results are presented in Figure 6.13. The figure shows the results from the colour shading experiment, sinusoids only (at 0.04 fixed luminance contrast), in comparison to the results from the colour matching experiments (raw data) at target value 0.075 for the two colour masks plus the baseline (colour mask contrast = 0).

The statistical analyses of both experiments were kept from the original chapters. The colour shading experiment data were tested using a Kruskal-Wallis test and when the depth modulation (by colour contrast) was significant the data is shown with accompanying fit (solid line). The contrast matching data analysis involved a t-test of each mask condition against the baseline (see Section 6.2.2 for more details); these test results are presented in Figure 6.13, for both chromatic masks against the baseline. Note that the data points for the depth experiments (filled circles) show the average of 10 data points, whereas the match data show the average of 40 data points (averaged across 4 sessions).

First, I will look into data from participants with significant effects in the depth experiment. Unfortunately, I could only retest two participants with significant effects. Recall that from the first experiment (depth matching), few participants showed a significant modulation of depth by chromatic contrast. However, I managed to have two participants, one with increasing and one with decreasing depth, M14/P5 and M19/P4, respectively. Therefore, it is possible to test the hypothesis relating colour shading to contrast interactions formally. One participant (M14) should show an increase in perceived contrast and the other one (M19) should demonstrate a decrease according to the hypothesis formulated earlier on. The hypothesis was that the changes in shape-from-shading are due to changes in luminance contrast perception; hence, an increased depth percept would have been due to an increase in luminance contrast and a decrease due to a reduced luminance contrast.

Looking at the results in Figure 6.13, this is not the case. It is intriguing, however, that observer M14/P5 shows strong effects in both experiments, however going in the opposite direction of what I expected. Furthermore, participant M19/P4 shows the expected decrease in both directions. What is remarkable here is that in both experiments the data plateau at small colour contrasts before receding at higher values (significantly in both experiments), which is the exact behaviour I would expect following the hypothesis formulated at the beginning of this chapter (and in the previous paragraph).

Now the reverse expectation is that for participants with non-significant effects in the

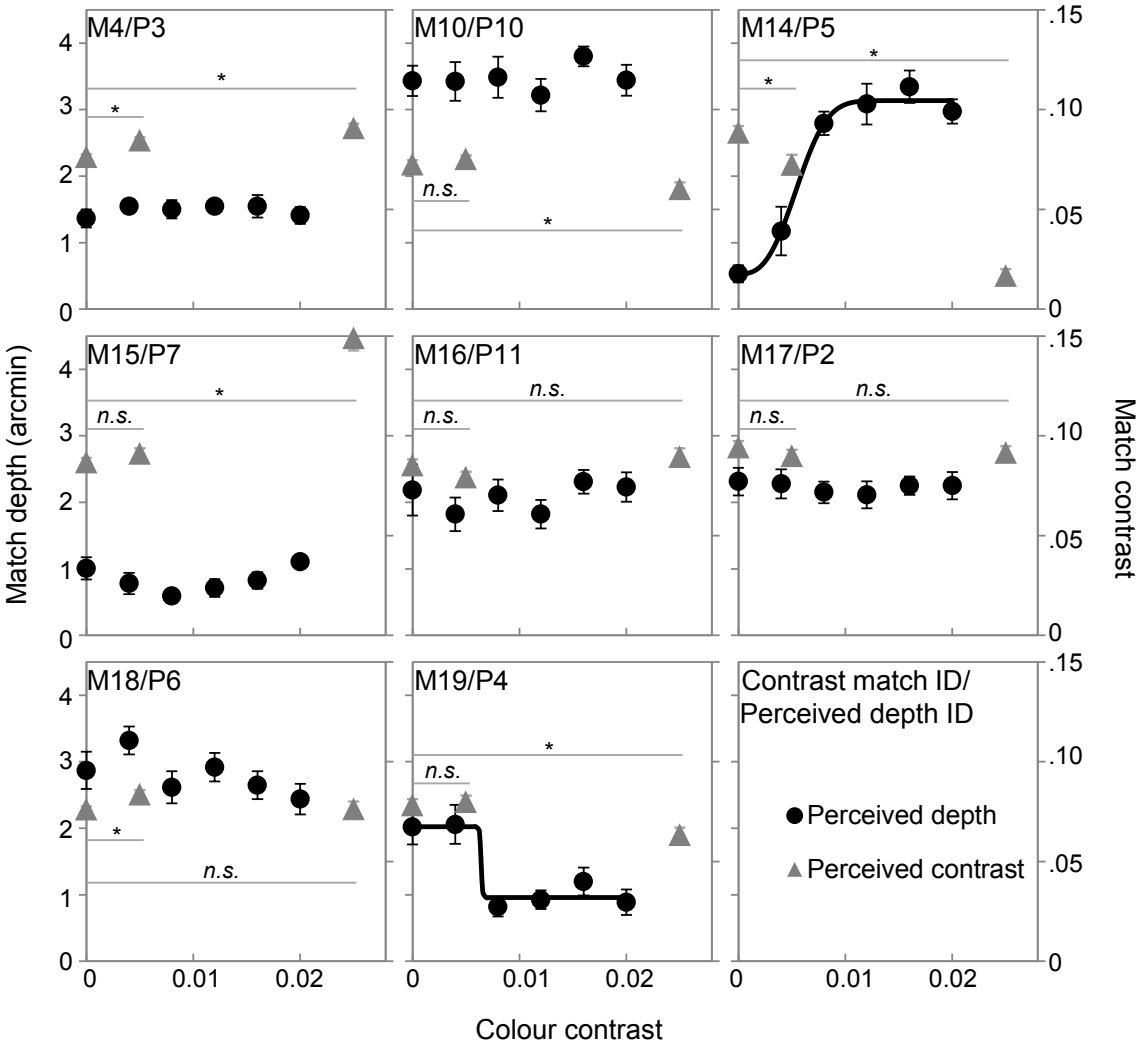


Figure 6.13: Comparison between depth matching and contrast matching experiments with a subset of results from participants who performed both experiments. The colour shading results (black circles) are presented as perceived matched depth of corrugation (arcmin, left axis) for a luminance component of 0.04. When there was a significant effect of contrast, the data are presented with a fit (see Chapter 5 for details). The contrast matching results (grey triangles) are presented as match contrast (Michelson contrast, right axis); the target value was 0.075. The baseline match is presented with a zero contrast chromatic component; the two other conditions are tested against the baseline.

depth matching experiment, there should be no difference in contrast matching (compare values shown in Figure 6.13). This is clearly the case for two participants (M16/P11 and M17/P2). However, participant M4/P3 is a counter example, this participant shows no modulation of perceived depth yet the perceived contrast was modulated; this is also the case for participant M18/P6. However, this participant (M18) has less clear data. Finally, participants M10/P10 and M15/P7 show no significant effect for the colour mask within the range of the first experiment tested; however, they do show a significant effect for the mask outside this range (especially M15/P7). It is hazardous to predict values outside the range tested, i.e. extrapolation, so in this case the evidence could go for either hypothesis.

The results are not clear-cut and it is difficult to draw conclusions from this data alone. I stress again the fact that the luminance components were not exactly at the same value and the chromatic components were also tested at a different, but overlapping, extent. I will elaborate some of these points in the following Discussion section as well as framing a hypothesis of why participant M14/P5 shows this specific response pattern.

6.3 Discussion

6.3.1 Interactions between luminance and colour

The study of the interactions between luminance and colour is the core aim of this thesis. As I have detailed previously, if both suppressive and enhancing effects are found, three main relationships can be expected between the effect of luminance and colour masks. Namely, those relationships types are correlated, anti-correlated, and non-specific. Here, I am discussing results at the level of the population (of observers tested); I will detail individual differences later on.

Correlated or anti-correlated There have been numerous suggestions in the literature that actually the (psychophysical) chromatic & luminance channels are not that separated (e.g. Gur & Akri, 1992, plus see Physiology section on dual encoding, Section 2.2.3). The assumption in my experiments was that the luminance channel might be completely achromatic and not excited by isoluminant stimuli. However, there is no proof that the chromatic channel does not respond somewhat to luminance (Chen et al., 2000a; Gur & Akri, 1992; Regan, 1991)). There are three possible kinds of interactions happening. The first is an orthogonal chromatic modulation effect (XOM) on luminance, the second a luminance XOM on luminance, and finally a chromatic XOM on luminance effects due to colour stimuli input into a mainly luminance processing pathway.

A correlated effect can be seen as the easiest to interpret; a similar effect tends to be associated with a similar mechanism. In this dataset, if the population behaviour were correlated, then most effects would be suppressive/suppressive effects or enhancing/enhancing effects for the luminance/chromatic pairs. The implication in terms of heuristics is that the observer's visual system would be using a unique *rule* for orthogonal

processes, whether it is suppressing or enhancing. However, there is no consistency in the population on how to treat contrast interaction at suprathreshold, as the data does not fall in the enhancing or suppressing quadrants only (Figure 6.9). Furthermore, I found that the data did not support this hypothesis; the small correlation observed is largely driven by one participant with high suppression in both conditions.

One could argue that if there were to be luminance artifacts, correlated effects would be expected. This is however no definite proof. Such results could be due to similar mechanisms, heuristics (strategy), as stated before, and should be treated on a case-by-case basis, as the pattern of correlation is not found in all participants but on the population as a whole. As stated earlier, the correlation on the whole population is mild and dependant on target contrast. Thus it seems safe to assume there were no luminance artifacts.

The anti-correlated case is possibly the most interesting; however, it does not seem to represent most of the population's behaviour. Anti-correlation suggests differential treatment of luminance and colour in spatial vision processing of orthogonal luminance targets. Concretely, in the data, the luminance might have an increasing effect of perceived contrast and the colour component a decreasing one. The fact that this was not found can tell us more about the nature of those interactions.

Differential treatment would have suggested the presence of heuristics that treat the two signals differently and, hence, could extract meaningful information from it (Rubin & Richards, 1982). In suggesting a set of heuristics, or *bag of tricks* (as described by Ramachandran, 1991 and discussed in Kingdom, 2008), one should consider their effect on low-level vision, especially if that is the level at which these are implemented. Having different rules to process the colour and luminance interactions with luminance targets is consistent with anti-correlated effects in the data.

Furthermore, there is a popular idea that the visual system tries to eliminate redundancy in the signal as much as possible (e.g. Hansen & Gegenfurtner, 2009). However, we must consider when the information is combined again and how features are bound. As previously mentioned, the data overall do not follow this possibly interesting pattern.

Non-specific The other major alternative is that the relationship between the two processing streams is non-specific (at the population level). This might suggest two things regarding these interactions.

Firstly, I consider that these heuristics or processing interactions, depending on the level at which we want to interpret them, might be fixed and part of the common physiological connections of chromatic and luminance encoding neurons. In this case a non-specific relationship might suggest that the orthogonal interactions mediated by colour are a by-product of binding and that the modulatory effects might not serve a physiological processing role, or at least the role associated with reflectance versus light extraction. This also suggests that the binding is done differently in different brains and it brings no new processing or a faithful binding of features.

Secondly, I consider the possibility of learned (non-fixed) inter-modulations, i.e. connection formation in the visual system. Formulated differently, this is the possibility that orthogonal connections can be adaptively learned depending on the statistics of the visual world. If these interactions are based on learned statistics of the visual environment, then we could expect our participants to have had different visual experiences during the period these connections or heuristics were formed.

Alternatively, consider the possibility that actually these rules can be learned quite fast (see Dovencioglu, Welchman, & Schofield, 2013, for an interesting parallel regarding the learning of second order luminance cue interactions). If the learning is fast then recent visual experience could modulate the way we perform processing in spatial vision. This makes sense in terms of a dynamical environment in which observers need to adapt quickly. These fast adaptations of colour-luminance interactions are speculative and would need further inspection if this hypothesis were deemed more plausible than the others mentioned above.

An interesting step would be to start from the visual world statistics and build up predictions. As mentioned earlier, the work by Hansen and Gegenfurtner (2009) suggested independence between luminance and chromatic edges. It would be interesting to focus on the relationship between adjacent orthogonal edges, in both the luminance to luminance and the luminance to chromatic case. One might expect these to be completely independent and then it should be of no surprise that on average in the population no specific relation is learned between the two. A possible way to test this hypothesis would be to change the statistics to vary the relationship between edges and see if this influences the processing. This would require prolonged exposure and should be disassociated from adaptation effects. However, one might argue that the role of adaptation is to adapt to these dynamical changes in statistics; here, the temporality characteristics of these effects in terms of stimulus exposure and time duration after exposure would be crucial.

Summary The data tend to suggest that in the population, depending on the target level, there is a non-specific to slightly correlated relationship between chromatic and luminance suprathreshold masking. This could suggest that these interactions do not play a crucial functional role or that the statistics of real world images tend not to show specific relationships when considering luminance versus luminance and luminance versus red-green edges (both orthogonal). This would, however, require more research (on image statistics and the dynamics of adaptation) to demonstrate.

6.3.2 Individual differences

The individual variations in behaviour are a major part of the results found in this experiment. I will link this to several points mentioned in the previous chapter on colour-shading (Chapter 5).

In the colour-shading data chapter, I was assessing the existence and generality of processing heuristics of spatial vision, suggested by several experiments from Kingdom's

lab. From the variety of results, two main hypotheses were formulated: the existence of low-level interactions responsible for these effects, and the possibility that the heuristics used by observers are different.[8] Here, this experiment might be considered as *middle level* as suprathreshold contrast perception might already involve higher processes related to scene understanding (see Introduction on scene interpretation 3.6).

It is clear that the participants showed different behaviours; however, as was concluded in the previous chapter, this does not preclude the participants' responses to be based on heuristics. This is actually tricky to disprove. The interpretation of the results based on the heuristics assumption is that participants are using different ones, possibly dependent on their visual experience or the task in itself. Alternatively, it is possible that these effects are due to individual differences in channel connections that are not due to reflectance extractions. I will discuss the relation between the colour-shading experiment and the contrast-matching experiment in the following section.

6.3.3 Link to colour-shading

The analysis' results regarding the colour-shading effect and this data were not as clear-cut as anticipated. It is difficult to conclude if low-level masking is responsible for the observed effects. There is no strong evidence for the suggested hypothesis.

However, participant M14/P5 is particularly interesting as his behaviour is showing strong effects in both perceived contrast and perceived depth. For this participant, neither effect followed the expected pattern, perceived contrast decreased while perceived depth increased, both drastically (see Figure 6.13).

In the low-level explanation of the colour-shading effect I considered a purely feed-forward effect; however, as depth-from-shading gets extracted, it is possible that these processes feedback to low-level processing. Namely, what is accounted for by shape should not be accounted for in terms of reflectance. Then, if some of the intensity variation is explained by one factor, the accounted-for variation should be taken out of the other percept. This is reminiscent of the separated modules in Adelson and Pentland (1996), where each module extracts one effect (e.g., lighting of the scene) but each module can inform the other ones. Therefore, strong shape-from-shading perception should be linked to stronger suppressing effects in the luminance matching task. The logical conclusion is that this participant was experiencing shape-from-shading effects during the contrast matching experiment.[9]

6.3.4 Effect of colour

The results demonstrated enhancing and suppressing effects on orthogonal luminance suprathreshold contrast, as well as mixed effects for some participants. This has some

[8] These two hypotheses are in fact not mutually exclusive, as nowhere in the literature, to the extent of my knowledge, it is suggested how these heuristics are implemented.

[9] Unfortunately, I did not ask about depth perception in the debriefing of this experiment and, therefore, I cannot test the hypothesis.

implications for the processing of natural scenes. It suggests that luminance is not perceived directly, through the output of the LGN, but at a higher level when luminance and chromatic (i.e. red-green chromatic) information are combined. It seems that, as opposed to detection were luminance and chromatic are independent (see Section 3.4.4), suprathreshold perception of luminance is not independent. The perceived contrast can be modulated consistently; the role of the modulation, if any, is a broader question.

Consequently, when colour and luminance stimuli are presented together, even if the luminance is controlled properly by the experimenter, there will be some perceptual differences between observers depending on the cross-modulation effects. To my knowledge, this is a new finding in the literature on colour and spatial vision.

This might be compared to the *Helmholtz–Kohlrausch effect* (H-K effect), where chromatic strength influences perceived brightness for equal luminance[10] (Nayatani, 1998), or formulated differently, there is an increase in brightness with increased colour saturation. However, our stimuli were composed of two spatially modulated components, and not one test patch. None of the current models can fully account for the H-K effect (Withouck et al., 2013). Sanchez and Fairchild (2002) showed that with 22 participants the distribution of the effects varied significantly (in Sanchez & Fairchild, 2002, see Figure 4 distribution of effects, for red and green patches). However, with plaid patterns, instead of a uniform shape against the background, it is difficult to compare the two effects. Furthermore, sinusoid and plaid stimuli interactions might suggest a cortical site, whereas the source of the H-K effect is unknown.

Furthermore, my result regarding colour effects, could suggest that colour spaces that consider luminance and colour to be orthogonal are not entirely accurate in the suprathreshold domain. It would be of interest to study colour perception using a striped pattern with added luminance, to obtain monochromatic colour-matching functions and luminance efficiency.

It is known that the spatial content of a scene can interact with the colour percept (e.g. Werner, 2003) but interestingly our data shows that the chromatic information might change the spatial percept in more ways than are currently known (see Chapter 3). This is an interesting question for further colour science research. Furthermore, the suprathreshold contrast perception is under-represented in the literature; when suprathreshold stimuli are used, the tasks usually involve detecting higher-level properties (such as object features). The variability of effects between participants raises two types of question: first, the reliability of the testing; and second, if the experiments are correctly performed, the source of this variability in individual visual systems. These issues will be addressed in a following section (see Section 6.3.6, Limitation).

[10] See Chapter 3 on brightness/luminance.

6.3.5 Luminance

Both suppressing and enhancing effects were found for the luminance orthogonal masks. Once again, this has some implications for scene perception. With more complex scenes, and intricate combinations of edges at different relative angles and contrasts, it would become more difficult to predict the perceived contrast. This issue of pooling between channels, such as orientation, SF, and even chromatic and luminance, should be addressed as well as pooling across space. The data do not seem to show a categorical break in luminance contrast perception when tested across our range of luminance tests from 0 to 0.5 contrast (Brannan & Bodis-Wollner, 1991).

Interaction of luminance channels Campbell and Robson (1968) demonstrated independence of luminance channels at detection levels by testing square-wave and sinusoid detection. The detection was based only on the highest harmonics, no pooling or summation between channels was apparent.

It has been shown that the contrast of a square-wave pattern is perceived as higher than the contrast of a sinusoid (Ginsburg et al., 1980), which is evidence of pooling of contrast over several spatial channels. However, as noted by Hamerly, Quick, and Reichert (1977), at suprathreshold levels it is possible that the transducer function becomes "an aggregate of channels". Quick, Hamerly, and Reichert (1976), one of the first to study the question at the time, noted that the known spatial processing mechanism at detection level, i.e. the separate channels hypothesis, is not certain to exist for suprathreshold perception. Quick et al. (1976), using summation of two gratings of different SF[11], found some summation where the two frequencies equally contribute. Quick et al. (1976) found different transducer exponents for low ($< 5\%$) and mid contrast ($5 - 20\%$), meaning the summation was less efficient at low contrast values. However, using square-wave and sinusoids, Ginsburg et al. (1980) found similar results to those at detection (Campbell & Robson, 1968) at values higher than 2% contrast[12] with the main harmonic carrying all contrast perception (i.e. no summation, winner takes all).

Georgeson and Shackleton (1994) studied the contrast perception of plaids (matching overall plaid contrast to one grating) and developed an additive model seemingly reconciling various results mentioned above, although their model specifically did not allow spatial pooling. Additionally, Georgeson and Meese (1997) found that adapting to an oriented component could change the appearance of a plaid stimulus from plaid (a unified percept) to separate components.

Spatial pooling Spatial pooling is often ignored or considered not to happen in the literature; however, it can tackle some broader issues such as binding, texture encoding, and population coding (Meese & Baker, 2011). Meese and Baker (2011) outlined two assumptions ("dogma") in the spatial vision literature that they believe is not backed up

[11] Frequency of components: f+3f.

[12] Ginsburg et al. (1980) did mention the discrepancy with Hamerly et al.'s (1977) data.

by enough evidence. The first one is the probabilistic nature of spatial pooling, where the increased area of spatial pooling is a merely increased chance for one channel to pick up the signal (Legge & Foley, 1980, e.g.). The second one is that spatial pooling only exists at threshold levels.

Meese and Summers (2007) and Meese and Baker (2011) tested spatial pooling effects (with a detection task) across a wide range of mask pedestal values using a "Swiss cheese" stimulus, or double checker board, where the holes are uniquely present in one eye forming a uniform percept through dichoptic vision. It was found that spatial pooling was present and performed between the eyes and at all pedestal contrast levels. In other words, strong binocular effects (summation) were found.

They modelled their results with a three-stage model: the first stage was a binocular summation of similarly oriented filters, the second stage was spatial pooling[13] (in now binocular space), and the third stage was a gain control stage of the previous output. The addition of differently oriented filters in the second stage would permit extension of the model to cross-interaction effects. The addition of colour might be a bit more complex and might require combining other models. Meese and Baker (2011) adjusted the second stage so that the divisive inhibition is modulated differently with excitation input and area, as sort of suppression of inhibition. There is an interesting parallel here with colour-luminance effects as Vimal (1998a) suggested that luminance and colour segregation at threshold is followed by a common divisive pool at suprathreshold levels, and the facilitation could be due to a suppression of inhibition (disinhibition) between channels. This 'cross disinhibition' between channels might be happening in some cases in our results and explain some of the facilitatory effects. Spatial pooling might also play a role and it would be interesting to vary the size of the test and match stimuli to test the role of this mechanism.

A few comments on metrics This section refers back to the distinction made in the Methodology Chapter 4 on the distinctions between physical properties of the stimulus and hence physical metrics and perceptual metrics (such as the photopic curve for example). Furthermore, perceptual metrics are usually based on a *standard observer*, which is sometimes not accurate enough (see Chapter 8, on Equiluminance, for distinction between nominal isoluminance and observer defined isoluminance).

Coming back to complex scene metrics, it is possible that the full description of the stimuli is given by contrast metrics at all orientations, as one number (based on highest and lowest values) might be not sufficient to detail orientational effects. Peli (1990) argued for a contrast metric for each frequency band at any point of an image, and one could add orientation. However, this becomes complex when adding phase effects. In particular, regarding physiology data on surround suppression, one might expect to have local contrast modulated by flanking stimuli. Here, the distinction between contrast and population

[13]Interestingly, Meese and Baker (2011) described the spatial summation as an emergent property with no special role of dichoptic stimulation.

encoding of the whole scene is blurred. Spatial vision is often studied as a greyscale world, however, as we have seen from the previous section the introduction of colour adds another layer of processing that should not be ignored and should be added in scene contrast metrics.

Effects at low levels The low luminance masks in general did not have effects on the target's perception. This has some implications in terms of luminance artifacts in the colour stimuli. Assuming that small luminance artifacts are remaining after the equiluminance calibration, these luminance components should have negligible effects on perceived contrast. This can be seen in Figure 6.4. Furthermore, using the method of adjustment might not be precise enough to pick up small variations in contrast perceptions; which is why most of the analysis was performed with the higher mask contrast data. Other limitations will be discussed in the next section.

6.3.6 Limitations

My study used a matching paradigm with free viewing and the two patterns visible at the same time by the participant, similarly to Georgeson and Sullivan (1975), as opposed to a 2IFC method with timed presentation (e.g. Cannon & Fullenkamp, 1991b; Switkes & Crognale, 1999; Vimal, 2000). Quick et al. (1976) used a timed matching paradigm; observers had a time window of 4s and noted that it was ample time for observers to do so. They also noted that settings were consistent between sessions. The analysis in Section 6.2.3 seems to corroborate Quick et al.'s (1976) observations. I do not find adaptive effects during the experiments and although presentation time was not windowed (in my contrast matching experiment), participants tended to be fast in their matching. The average response time was in the vicinity of the times in Quick et al.'s (1976)'s paradigm.

Adaptation happens in the span of several seconds, and it is tuned to SF and orientation (Blakemore & Campbell, 1969; Blakemore, Muncey, & Ridley, 1973). Consequently, I do not expect cross-adaptation effects between mask and target. It would be an improvement on my design to use a timed presentation window and a long timed blank screen between trials. A 2IFC design might bring even more precision at the detriment of time and number of conditions tested. Furthermore, Georgeson and Meese (1997) showed that a short adaptation time[14] could change how a plaid is perceived depending on the contrast of the adaptive plaid. Nevertheless, their task specifically compared plaid versus component perception, whereas our task always involved retrieving the perceived contrast about one component. This would argue for the chromatic task to be easier than the luminance task, as the component contrast might be more easily separated when composed of luminance and colour rather than luminance-luminance. An interesting extension of our experiment would be to ask our participants about plaid perception, including depth, as perceptual feedback could change the contrast perception. Changes in the perception of the stimuli

[14] 5 minutes initial adaptation plus 10 sec between trials, stimulus presentation time 500 msec.

(and its components) could explain some of the variability of the dataset.

6.3.7 Choice of task and participants' strategy

In order to judge the amount of contrast perceived in the left-oblique direction, a matching left-oriented luminance stimulus with only one luminance component was used (see Method Section). The task was to extract the (luminance) contrast in one orientation and reproduce it in a similar looking matching stimulus. This is not a natural task and it assumes that the participant is able to extract information independently from different spatial channels (see Chapter 3 on Psychophysical channels). However, it is also part of the research question that I am trying to answer. If the spatial channels interact strongly (at suprathreshold), then it becomes more difficult to extract the signal from one channel without it being modulated by parallel channels processing other orientations or other information (i.e. chromatic). I am trying to obtain a measure of this interaction with the experiment presented in this chapter.

The assumption that I made is that participants are able to access and estimate the amount of energy or contrast in a spatial channel without having to reconstruct it through the use of a strategy to perform the task. I made sure to add cases with no orthogonal component (baseline cases) interleaved in the design so that the participants did not feel like they were switching task but were always performing the same orientation contrast estimation task.

At suprathreshold level, components can be fused and interpreted as surface or texture instead of the sum of individual spatial channels. This fits with the hierarchical view of vision in which low level stages process increasingly complex and specified patterns. There is also physiological hierarchy, the receptive fields of neurons become increasingly bigger and encode more complex features (from sharp On/Off light encoding in the retina to face encoding cells in the Inferior Temporal cortex for example). It is still unknown how the final visual percept is created and which levels of the visual hierarchy are responsible for perceptual consciousness. We can make the assumption that visual information at each processing stage is available. Note that potentially, interactions at higher levels could be seen at intermediate stages through feedback connections.

Nevertheless, one could argue that this assumption is wrong and as I did not ask specifically how my participant were performing the task, it is possible that they used specific strategies. In the rest of this section, I will discuss a particular strategy that could lead to some of the effects observed.

If the participants were seeing one unified textured surface, they could have resorted to reconstruct the individual components (or only the luminance component of interest) through the use of sampling within the stimulus. By sampling along the orientation of choice, they could afterward match those samples statistics with the contrast in the matching stimulus. Furthermore, in the case of luminance mask conditions, the sampling of luminance would be strongly affected by the two components if the observers sampled

preferentially at specific spots of the stimuli. Essentially, they would compute a sort of local contrast. This can be a problem for the luminance mask; however, this argument does not hold for the chromatic mask because there were made isoluminant (see Methods and Isoluminant chapters, Chapter 4 & 8). Even if a participant used a local sampling strategy for chromatic parts there is no added luminance to the stimulus (from the chromatic component) that could interfere with the percept of contrast. Consequently, the chromatic effects observed would still be due to channel interactions and not participant strategy. The same conclusion cannot be made for luminance masks.

6.3.8 Comparison to previous studies

Regarding luminance and colour interaction, Kingdom, Bell, Gheorghiu, and Malkoc (2010) used a stimulus composed of disks of luminance and iso-luminant colours using 2IFC "most salient" or "highest perceived contrast task". They generally found a masking effect happening, however, the measure of contrast was a combined one between luminance and chromatic (red-green or yellow-blue). Kingdom et al. (2010) admitted that the results could have several explanations. More research with this kind of stimulus is required. It would particularly be interesting to try combining the protocol used in this Chapter's experiment, contrast matched in one spatial orientation, with the kind of stimuli used by Kingdom et al. (2010). Purely, achromatic variations could be studied this way as well.

Compared to data by Meese and Hess (2004), who were the first to measure suprathreshold contrast perception of one component in a masking paradigm (luminance only), the data in this chapter matches the luminance binocular data with both enhancing and suppressing effects (Meese & Hess, 2004, Figure 5 & 6, inverse triangle symbols). Once again it would be interesting to incorporate some of their protocols (dichoptic and monocular). The enhancing contrast effect should be explored and not dismissed, as the data shows that it happened for both participants in Meese and Hess (2004) but at two different low frequencies. If this enhancing effect was due to strategy or an adaptation effect it should happen for the two spatial frequencies equally, as there is no reason why these two explanations would vary between 0.46 and 1 cpd. Regarding both my data and those from Meese and Hess (2004), the facilitatory effect appears dependant on binocular presentation, relative contrast (seen in this chapter), SF, and observer (Meese & Hess, 2004).

6.4 Summary

In this chapter, I have shown the novel finding that orthogonal chromatic masks do have an effect on the perceived contrast of suprathreshold luminance targets for some participants. The effect can be suppressive or enhancing depending on the participant, which is in-line with the results presented in the previous experimental chapter on the colour-shading effect. However, comparative results with participants having performed both tasks failed to reveal a conclusive systematic link between the two.

Additionally, I have shown that similar suppressive and enhancing effects can be found using luminance masks. This challenges the usual suppressive nature of orthogonal components that is assumed in the literature due to a focus on detection (e.g. Petrov et al., 2005). Once again, large individual differences were found within the group of participants tested. The similarity between chromatic and luminance masks was analysed and discussed. These new results might be important to understand how scene characteristics are processed in the brain.

Chapter 7

Orthogonal masking study at detection level

7.1 Introduction

7.1.1 Aim

The aim of this chapter is to explore the orthogonal modulatory effects of colour/luminance masks, i.e. suppressive or enhancing, that were found in the previous chapters 5 & 6, but this time at the detection level. In this chapter, I will be considering the detection of luminance targets and measuring how luminance and red-green chromatic masks are interacting with luminance processing.

In the previous chapter, the data showed influence of the chromatic information on the perception of contrast of luminance targets. If the mechanisms responsible for the phenomenon are present at a low level processing stage it might be possible to track them also at the detection level. There might be some specific role for the influence of colour on luminance perception and that could be found also at the detection level.

Additionally, the use of luminance masks in this experiment is again important in order to check for luminance artifacts, but also is of interest in itself, given the interesting results we obtained in the previous experiment.

7.1.2 Chapter overview

The literature on channel interactions (see Chapter 3) suggests that there is no modulatory interaction effect of the chromatic component on luminance detection. However, the best evidence of this was tested using iso-oriented (chromatic) masks and (luminance) targets (see Chen et al., 2000a). The results from the previous chapter suggest that interactions do exist between luminance and chromatic channels at suprathreshold level in an orthogonal configuration. In this chapter, I will try to find evidence for these interactions at the detection threshold. To this effect, I will present an experiment based on Chen et al.'s (2000a) design but modified to follow the same spatial characteristics used in the previous

chapter.

The experiment presented in this chapter focuses on orthogonal interactions; it aims at comparing luminance and colour orthogonal effects on a luminance target. This is presented in the Results Section 7.2, I then compare the results obtained from this experiment with the previous data on colour shading and suprathreshold masking and discuss the implications.

Usually in the literature, no facilitation interactions are found between luminance and chromatic components at low mask contrast (Regan, 1991), and only suppression is found at higher colour contrast. However, to the best of my knowledge, orthogonal interactions have never been tested for detection tasks in the context of cross-channel (here luminance versus red-green chromatic) interactions.

Here, as I am dealing with a detection task, some methodological adjustments and extra constraints were added. The main constraint was the amount of testing required, which in turn limits the number of participant that could be tested. The second constraint is in term of the stimulus; as it is a detection task the timing needs to be restrained (compared to the previous experiments). This means adding a temporal envelope to the stimulus contrast. This will be detailed in the stimulus methods. The experiment described here is based mainly on the protocol described in Chen et al. (2000a) with the exception of some key details, such as orientation. The specifics of the methodology I used are detailed in Section 7.1.4.

7.1.3 Hypothesis

Firstly, for the luminance stimuli, I am expecting to see the usual *dipper function* pattern, consisting of a facilitatory dip at low contrast values, followed by masking (i.e. suppressing effect). Facilitation was found, by Meese and Holmes (2007), mostly for stimuli of low SF and low TF. See Chapter 3 for review on masking at detection level.

In the colour masks case, there are several hypotheses that can be formulated. The conservative approach regarding the previous literature especially (e.g. Chen et al., 2000a) is that there will be no facilitatory effect at low level and masking occurring when contrast gets high. However if the effects obtained in previous chapters are similar at detection threshold we might expect to observe more complex patterns of effect which would be something not previously found in the literature.

7.1.4 Methods

Observers

In detection masking studies, such as this one, the number of sessions needed tends to limit the number of participants, for practical reasons. Therefore, this experiment has fewer participants than the previous ones. A total of 4 observers were tested. All observers were

Table 7.1: Participants ID tested in both detection and contrast matching experiments

Experiment	Participant ID			
Detection exp.	D1	D2	D3	D4
Matching experiment	M1	M6	M3	n.a.

naïve to the task and hypothesis tested with the exception of the author (D1)[1]. All of them had no colour deficiency as tested by the Ishihara colour plate test (38-plate edition, 1979). As for the previous experiments, Ethical approval was given by the University of St. Andrews ethics committee UTREC (reference PS6135) and testing followed university guidelines. Participants were given monetary compensation for their time.

The subject pool for this experiment includes two participants having performed both contrast-matching and detection experiments,[2] participant M3/D3 and D2/M6, as well as the author (D1/M1) as described in Table 7.1.

Apparatus

The setup is the same as for the previous experiment on luminance contrast-matching. The same CRT monitor was used (with consequently identical chromaticity coordinates). The guns' chromaticity values were obtained using the same equipment and methodology defined in Chapters 4 and 8, this goes as well for the gamma correction procedure. The viewing distance was identical (i.e. 96 cm).

General method

The experiment was a temporal two-interval forced choice detection task (2IFC). For each trial two disks were presented, in two different time intervals. In one interval, a luminance target was present and in the other one, it was absent.

In a masking study, the stimuli are composed of a target stimulus, present in one interval, and a mask of varying contrast (depending on the trial/block) present in both intervals. The mask could be of two types red-green chromatic and luminance.

The purpose is to find the perceptual threshold, usually defined as a specific point in the fitted psychometric function, as a function of contrast masks values (or TvC).

For each mask contrast tested, the contrast of the targets is varied to span the values from 0% detection to 100% detection, the threshold is then fitted and an estimate is obtained. The results are presented as a TvC curve with respective uncertainty error bars on each estimate.

[1]The nomenclature used in this Chapter to distinguish the participants from Experiment I and II is D1-D4, the author is D1. This code was used to distinguish participants from the depth matching experiment labelled P1 to P12 and contrast-matching experiment labelled M1 to M22.

[2]Note that none of the participants performed all three main experiments, and respective equiluminant setting experiments (see Chapter 8 in this thesis), however there are some overlaps between the groups, as described before.

Target Stimuli and Mask

Stimulus specifications The target stimulus was always a left-oriented luminance defined sinusoid. The mask was always a right oriented sinusoid. This is a major difference from the work by Chen et al. (2000a), whose procedure my experiment is based on, as they used a pedestal protocol.

The *baseline condition* in this case is when the mask contrast is zero, this condition can be used to express further contrast values (e.g. mask) in multiples of detection threshold. Note that the baseline conditions of the chromatic component (in this case right-oriented) were also tested. The mask was either chromatic or achromatic depending on the condition. Both masks and targets were sinusoids (as defined in Methods Chapter 4).

All sinusoids had the same spatial frequency (0.75 cpd), similarly to the previous experiment presented in Chapter 5. Chen et al. (2000a) used Gabor stimuli of similar SF (1 cpd). As with my previous experiments, mask and target were displayed on alternate frames and interleaved (see General Method Chapter 4). The size of stimuli was 4°; the edges were smoothed by using a Gaussian filter. I described stimuli generation and colour space in a previous Chapter 5.

Temporal waveform The stimuli were modulated in time using a truncated-Gaussian waveform (see Figure 7.1). The definition of the temporal waveform is given by the following set of equations as a function of time (t):

$$f : t \mapsto = \begin{cases} \frac{\mathcal{N}_{\mu,\sigma}(t)}{\mathcal{N}_{\mu,\sigma}(\mu)} & \text{if } t \in [\mu - \frac{w}{2}, \mu + \frac{w}{2}] \\ 0 & \text{otherwise} \end{cases} \tag{7.1}$$

This equation in a modified version of a Gaussian distribution $\mathcal{N}$, with μ, when $t = \mu$ the functions reaches its highest point, σ is the parameter changing the width of the waveform (based on the variance of the modified Gaussian), w is the temporal window beyond which the Gaussian is truncated. The highest point of the Gaussian distribution (i.e. $\mathcal{N}_{\mu,\sigma}(\mu)$) is used to normalize the function itself.[3]

The parameters for my experiment were based on the ones used in Chen et al. (2000a) and were as follows, width of the distribution $\sigma = 40ms$ (called scale parameter in Chen et al., 2000a), for convenience $\mu = 0$, 0 being the half-duration of an interval. The total duration of one interval was 1 s. The truncated window parameter was set as $w = 160ms$. The results of this function can be seen in Figure 7.1, showed by the continuous line.

The temporal envelope (Figure 7.1, continuous line) works by multiplying the cone contrast (see Methodology Equation 4.7, 4.8 & 4.9) for the three classes of cones by this Gaussian waveform. Beyond the truncated ends of the waveform the function is at floor level i.e. 0, consequently the cone contrast becomes zero. As I have described in the

[3]This modification, itself, notwithstanding the fact that it is truncated afterwards, makes the function not integrate to 1 anymore, and it is therefore not a probability function. However, this was not the purpose; this function was created to make a temporal *pulse*.

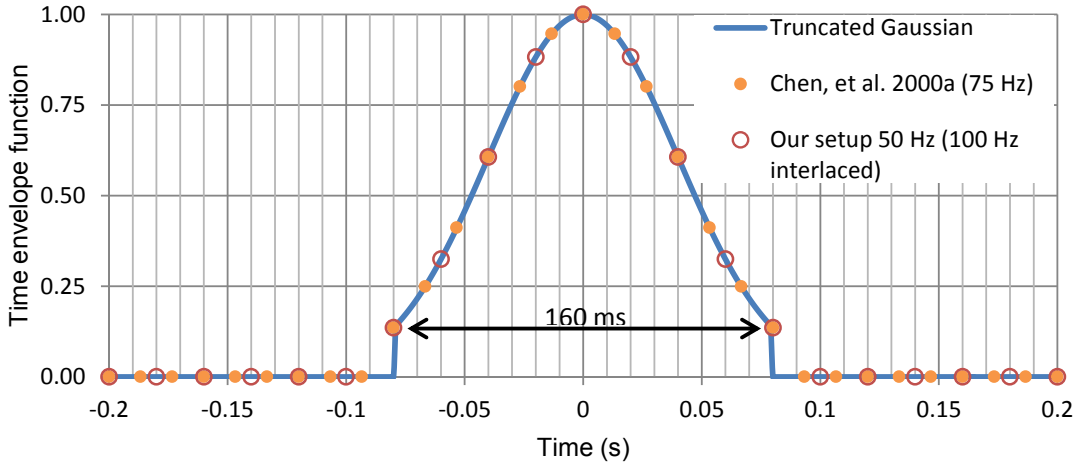


Figure 7.1: Temporal waveform of the stimuli as defined by Equation 7.1 with parameters, the contrast frames are shown for both the setup used in this experiment and the one by Chen et al. (2000a).

Methodology Chapter 4, null cone contrast corresponds to the background value. In Chen et al. (2000a) the background was a white point (CIE xyY: 0.28, 0.31, 29.6 $cd.m^2$) in this experiment the background was similar to the previous experiments presented in this thesis, i.e. grey (CIE xyY:0.282, 0.311, 40 $cd.m^{-2}$ similar to Kingdom et al., 2005).

Chen et al. (2000a) did not used interleaved mask and target frames as I did and so were displaying at 75 Hz directly, in my method as the two types of frames are interleaved (see Method) I consider pairs of frames together as if the refresh rate was halved. This was done so that both mask and target reach the highest point of the temporal waveform, in order to make the contrast reported accurate. Single frames from Chen et al. (2000a) (closed circles) and interleaved pairs of frames (open symbols) are shown in Figure 7.1.

Equiluminant setting

The equiluminant point was obtained using a hetero-flicker photometry (Anstis & Cavanagh, 1983) technique. The details of experiments are the same as the one for the contrast-matching equiluminance setting experiment. The participants performed one setup session before the experiment.

Blocks design

The design of this masking experiment was split in blocks. The number of block performed varied between participants. During a block the mask condition would be the same. Each testing session participants performed 2 to 3 pseudo-randomized blocks. Each block consisted of 100 trials of the same mask condition, reducing the uncertainty of trials of the same block. Participants could rest between blocks and trials, as they had to press the response button to trigger the next trial.

For each contrast value tested each participant performed 6 blocks or more. Partici-

Table 7.2: Details of orthogonal masks conditions tested (types and contrast) for each participant, B: baseline condition

	Contrast condition														
Part.	B	Colour masks					Luminance masks								
D1	0	.005	.01	.02	.03	.04	.01	.02	.03	.04	.05	.07	.10	.20	.40
D2	0	.005	.01	.02	.03	.04	.01	.03	.04	.07	.10	.20	.30	.40	n/a
D3	0	.01	.02	.04	n/a	n/a	.01	.02	.03	.06	.10	.40	n/a	n/a	n/a
D4	0	.01	.03	.04	n/a	n/a	.01	.02	.03	.04	.10	.30	.40	n/a	n/a

pants D3 and D4 performed 6 blocks per condition (60 and 66 respectively), participant D2 performed 6 block per condition for the luminance masks and 7 for the chromatic masks (total 89 blocks). Finally participant D1 (the author) performed 8 block per condition (total of 120 blocks).

The length of a block was 100 trials. During each block an adaptive staircase (see Section 7.1.4, below) selected the target conditions, (some of the author, ID D1, sessions at the beginning of the experiment lasted for 125 sessions instead of 100; the length was adjusted to have two to three sessions performed in one hour and sufficient number of trials to get a good estimate). Due to the greater number and length of sessions tested, the data for observer D1 is expected to be more accurate.[4]

Mask values tested were different between participants and selected by the author in order to find a potential dip (see Chapter 3, Figure 3.9) and then increase slope whose presence on the x-axis (masks contrast value) might vary depending on their detection threshold. The mask conditions tested, with luminance targets are detailed in Table 7.2 for all participants.

Staircase design

2IFC Bayesian staircase During each block, participants performed a 2IFC tasks with the same mask conditions throughout. I did not choose to use the method of constant stimuli but instead chose an adaptive methodology. The method used was a Bayesian staircase, I implemented (using Matlab) the Bayesian staircase described by Kontsevich and Tyler (1999). The main advantage of adaptive methods (such as here with the staircase) is that they adjust the values of contrast tested depending on the answers given by the participants during the session. This is a major advantage compared to the method of constant stimuli.

This method also relies on the knowledge of Psychophysical detection task; we already have a good idea of the results that we are going to obtain, namely a sigmoidal psychometric function from 50% to 100%, there is additional knowledge that can be used to make the procedure more efficient. For this experiment, I decided to use a Weibull function,[5]

[4]The total number of trials tested were, D1 (> 12000), D2 (8900), D3 (6000) & D4 (6600).

[5]In their original paper Kontsevich and Tyler (1999) used a cumulative Gaussian; however, the actual

see Chapter 4.

The main steps of Bayesian methods proposed by Kontsevich and Tyler (1999) are described below (see original paper for detailed explanations and mathematical formulations).

Implementation We need first to specify a psychometric function (here Weibull) and a probability distribution of the parameters. The probability distribution of these parameters at the beginning of the block is flat and bounded to meaningful values. The proper choice of boundary can speed up the convergence of the data. To further speed up the process Kontsevich and Tyler (1999) recommend to pre-compute conditional probabilities of the psychometric function, $p(response|slope, threshold, contrast)$. Once again this depends on the discretization of the parameters and their boundary; choices need to be sensible and can be adjusted to improve computing performance. The discretization of the contrast values is especially important, as it will determine the values that are going to be picked during the experiment. There is no need and no benefit to have a discretization of contrast parameters higher than the potential resolution of the screen in terms of contrast. I chose a discretization of 0.001 for contrast.

The next step is to calculate the expected outcome (response) correct or incorrect for the whole parameter space (i.e. all threshold and slopes and all contrast values). This posterior probability is evaluated for both correct and incorrect outcomes. The entropy is then calculated, the calculation involves collapsing the posterior probabilities across slope and threshold (leaving only contrast). This is done for both detection success and detection failure. At this stage, we have two entropy functions one for detection success and one for failure, these functions are defined as a function of contrast.

The next step involves computing the expected entropy by multiplying each computed entropy by their respective prior distribution and summing the two. The point of minimum entropy on the obtained results is where the *information gain* will be greatest, irrelevant of outcome. This is how the next value of contrast to be tested is chosen. The use of information theory is aimed at increasing outcome knowledge from the next trial. This is what makes this procedure efficient.

The next step is an experimental step. The participant is presented with the selected trial; the posterior probability (computed from a previous step) of the specific outcome whether detection failure or success is obtained then replaces the prior probability. We can now return to the first step of the algorithm. The updating of the prior knowledge reflects the added knowledge obtained from the previous trial. This is the part of the procedure that makes it an adaptive and Bayesian approach.

There is no stopping criterion regarding the data, as opposed to other staircase methodology, the session is stopped after a certain number of trials. In their paper, Kontsevich and Tyler (1999) detail how the length of a staircase affect the outcome estimate. They showed that threshold estimates converged fast (in less than 50 trials), and that the method was

function does not matter much.

robust to blank trials. I selected 100 trials length to obtain clean data (after initial value of 125).

It would be possible to keep the prior distribution between sessions of the similar masks type and contrast; however, I decided to always start with a flat prior. Additionally the highest point on the posterior distribution gives us the best estimate of the parameters using this psychometric function and could be used as estimate of threshold and slope results, however this is not the final fitting methods used, this is described in the Analysis section below.

Summary This adaptive procedure gives fast estimates by combining knowledge about how the contrast value relates to the behavioural response through the psychometric function as well as using information theory and Bayesian statistics. The collected trials were stored and saved in whole session files. The posterior distributions were not used in between sessions of the same values. I did not use the estimates within one session as final estimates of threshold. The final parameters estimates of the psychometric function were fitted using another method described below. I used the Bayesian staircase mainly to sample the contrast targets meaningfully for maximum information gain and fast collection time. The data compilation and analysis methodology is provided below.

Data compilation and analysis

All the data points were compiled for each individual participant at each contrast value tested. Data were subsequently fitted using a Weibull function Equation 5.1, p. 97. The fit was bootstrapped 400 times to obtain confidence intervals on the threshold parameter. An example of this procedure is provided in Figure 7.2. In this example the fitted contrast threshold was 0.131 ($\pm$0.0005). Note that the cumulative Weibull function threshold parameter fits the 80% correct response when bounded from 0.5 to 1. It has been scaled and translated from the original described in Colour-shading chapter, Section 5.2.1. Consequently, from this fit, I obtain two parameters, one threshold (80% correct) and a slope parameters expressing how steep the curve is (higher is steeper). The results will be presented using these two parameters in Section 7.2.

The number of data points collected for each target contrast is illustrated by the size of the data points (Figure 7.2). It can be appreciated that the adaptive method of the Bayesian staircase, described in the previous section, favoured testing contrast values in the region around high slope. This was done to maximize the information about the threshold parameter.

The example provided is an example of good fit in which we obtain a tight uncertainty. We can clearly see that the adaptive method did not dwell too much in contrast values where the results were already saturated ($\approx$ 100%).

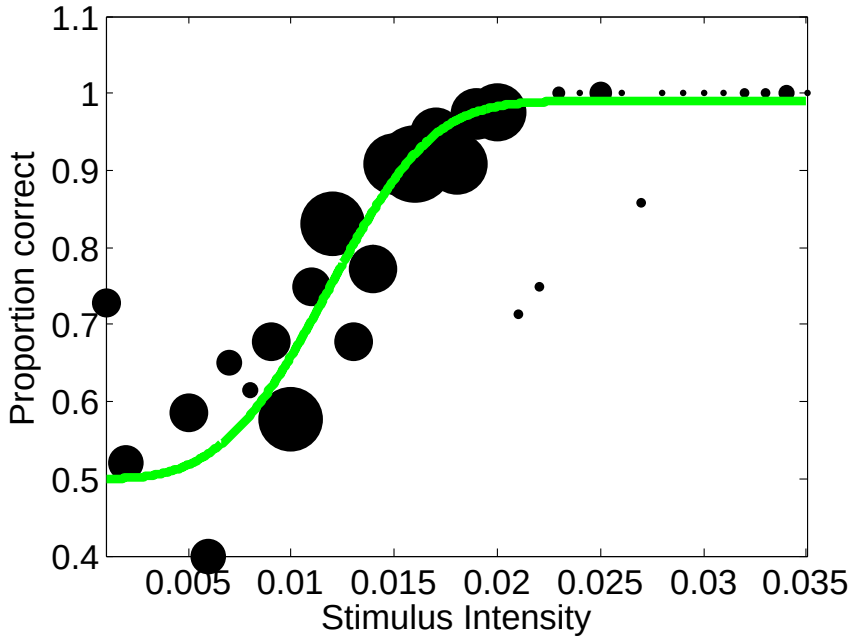


Figure 7.2: Example of psychometric function fitting procedure for one mask condition (10% luminance contrast) for participant D1/M1. The intensity of the stimulus corresponds to the contrast input of the stimulus. The size of the data points represent the number of trials tested. The fitted function is represented in solid green. The fitted threshold parameter (see text for method) is here 0.131 with bootstrapped interval of ± 0.0005.

7.2 Results

7.2.1 Luminance mask

The fitted threshold for each luminance masks conditions tested (see Table 7.2) following the method presented in Section 7.1.4 are presented in Figures 7.3, 7.4, 7.5 & 7.6. In each figure I present the threshold fitted and the bootstrapped standard error of the threshold parameter (Section 7.1.4).

The baseline, i.e. detection threshold without a mask, is presented in greyed area in both x-axis (mask contrast) and y-axis (luminance detection threshold). Any detection below the baseline is considered facilitation and detection threshold above correspond to suppressive masking. The classical dipper shape behaviour is facilitation at low contrast followed by masking at higher levels of contrast.

The dipper function pattern is visible, to a different extent, for all participants. However, in this dataset a different behaviour is emerging: at very low contrast there seems to be an initial masking before the facilitatory dip. The possible reason for this will be detailed in the Discussion section of this chapter. A brief explanation is that this detection at baseline might be improved by flicker detection and not pattern detection.

The continuous lines in Figure 7.3, 7.4, 7.5, & 7.6 represent the fits obtained based on a model by Meese and Holmes (2007). The model and fitting will be detailed in Section

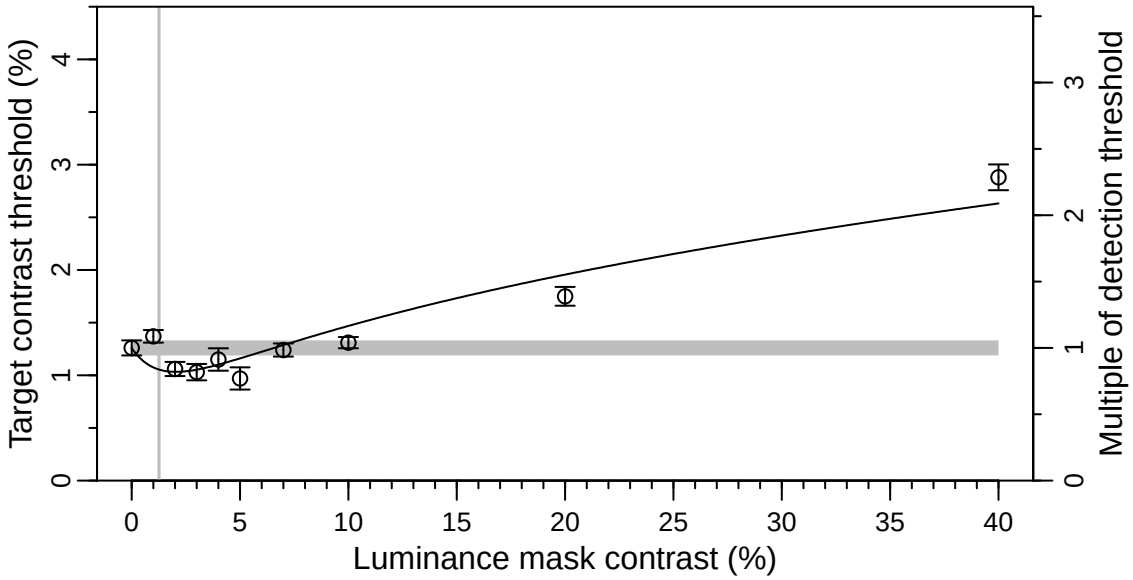


Figure 7.3: TvC response curve for luminance mask participant D1/M1. Grey area: detection threshold estimates for luminance without mask. Black line: dipper fit (see text).

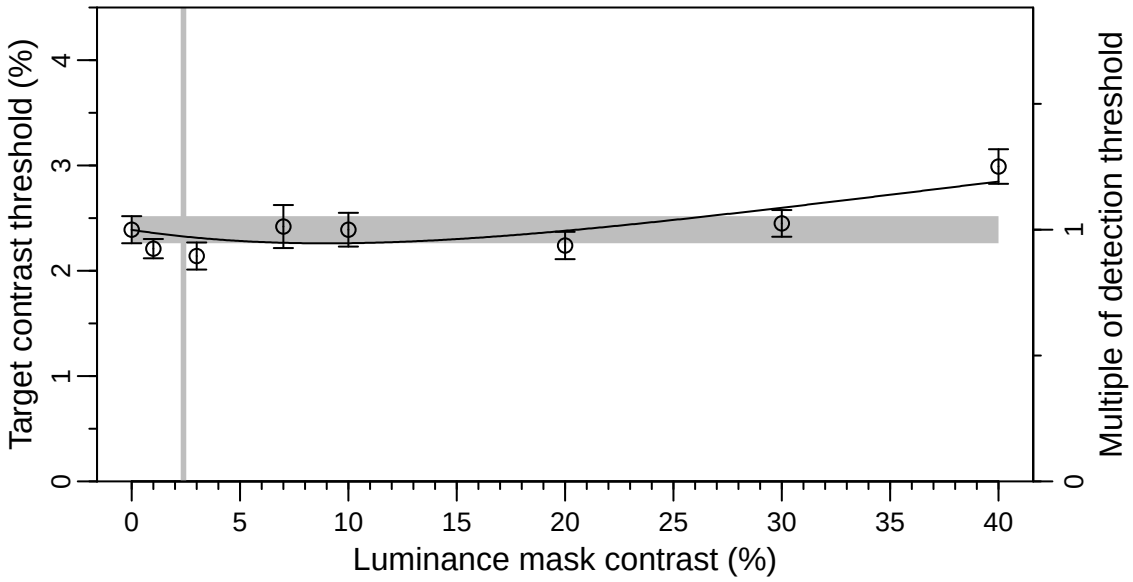


Figure 7.4: TvC response curve for luminance mask participant D2/M6. Grey area: detection threshold estimate for luminance without mask. Black line: dipper fit (see text).

7.2.3.

The main parameter fitted from the results was the threshold parameter. The second parameter, the slope parameter, requires more data to be accurate but it is interesting to look at the slope of the psychometric curve as it is an indicator of the noise in the detection system, and the slope has also been linked to uncertainty.

In Figures 7.7, 7.8, 7.9 & 7.10, I present the fitted slope parameter (with standard error

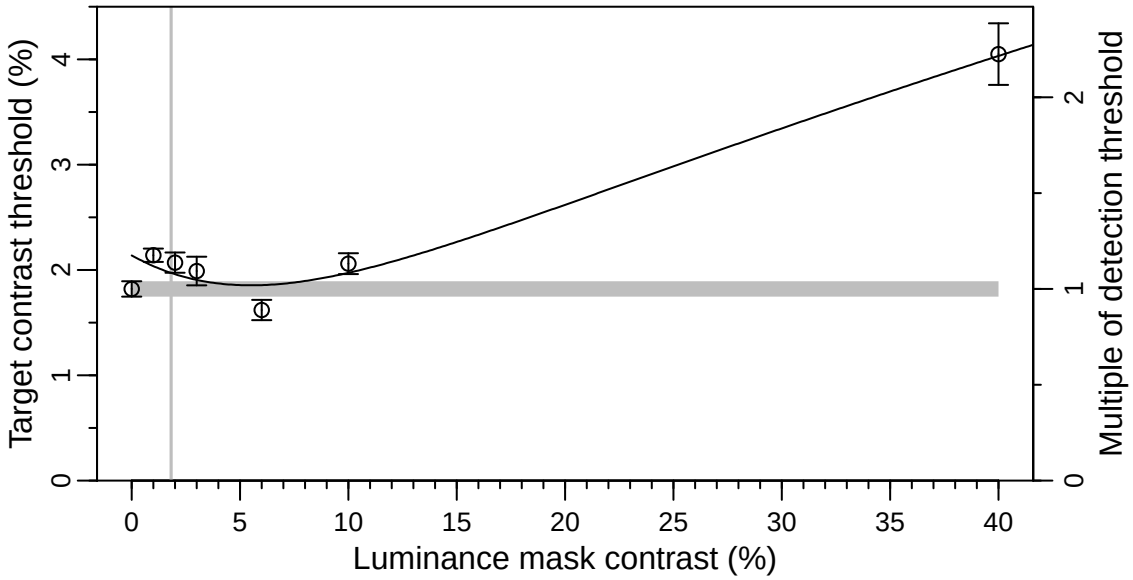


Figure 7.5: TvC response curve for luminance mask participant D3/M3. Grey area: detection threshold estimate for luminance without mask. Black line: dipper fit (see text).

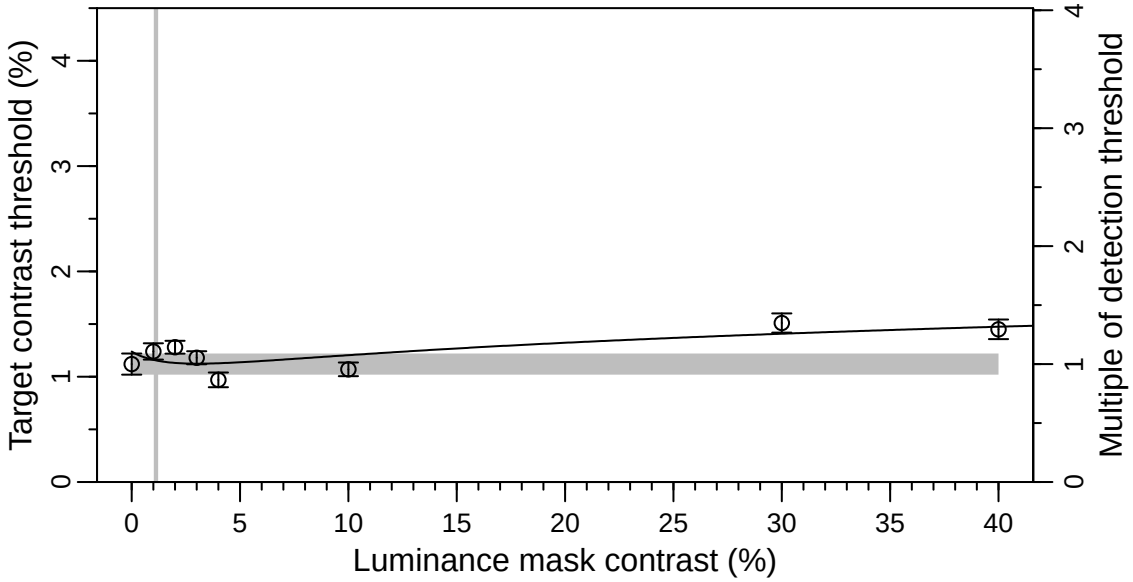


Figure 7.6: TvC response curve for luminance mask participant D4. Grey area: detection threshold estimate for luminance without mask. Black line: dipper fit (see text).

obtained from parametric bootstrap). In these figures, there is a tendency for the first data points to be higher than the rest, once again showing the specificity of the baseline condition. A high slope parameter corresponds to a steep or step-like function. Lower values of slope parameter correspond to shallower slopes. The baseline condition tends to be more step-like than the rest of the mask values. Additionally, it can be seen that mask contrast values with facilitatory effects in the thresholds tend to have shallower responses than the rest (see Figures 7.3, 7.4, 7.5 & 7.6). This contrasts with the highest mask contrast

values, which seems to lie in between the baseline condition and the facilitatory/shallower slope values.

Note that participant D2 did not seem to follow the patterns described above, the data (psychometric functions) from this participants tend to be noisier than the others are. This in turn can be seen with bigger parameters' uncertainty, hence the pattern of the data is more difficult to extract.

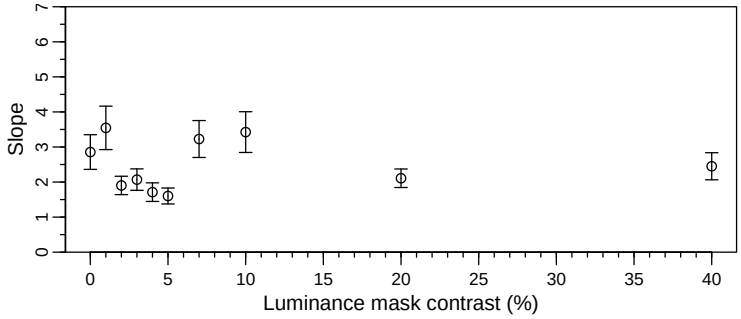


Figure 7.7: Fitted Slope parameter (Weibull function, see text) for the luminance mask, participant D1/M1.

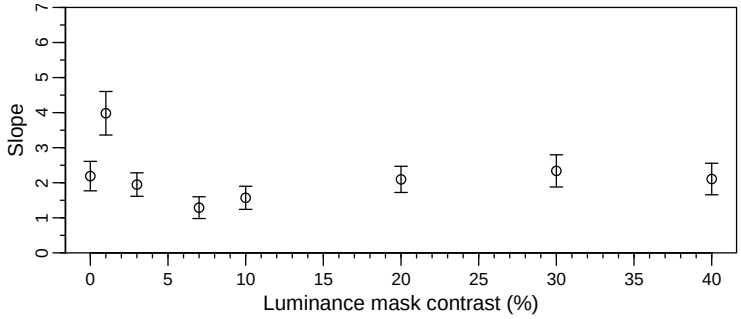


Figure 7.8: Fitted Slope parameter (Weibull function, see text) for the luminance mask, participant D2/M6.

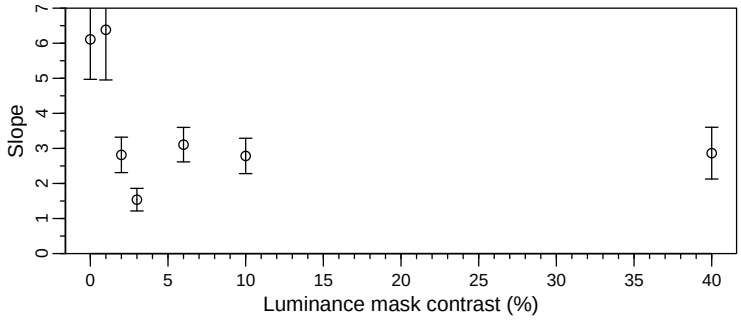


Figure 7.9: Fitted Slope parameter (Weibull function, see text) for the luminance mask, participant D3/M3.

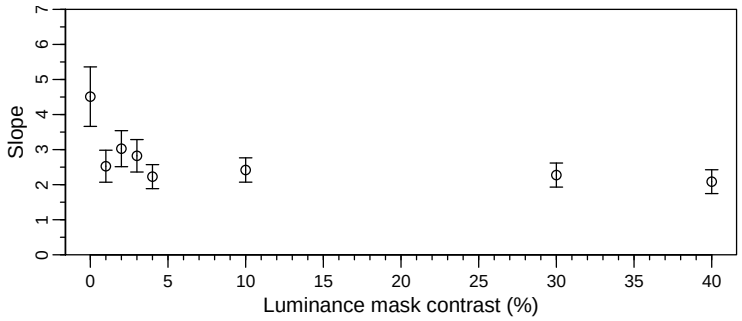


Figure 7.10: Fitted Slope parameter (Weibull function, see text) for the luminance mask, participant D4.

Individual differences

As described earlier, the amount of noise in the data is variable between participants. Additionally, participant D1 performed more sessions than the other participants did, which explains in part why the error bars are smaller.

Individual differences can be seen in the baseline condition with participant D4 having an exceptionally low threshold (see Figure 7.6). This participant also does not show a strong dipper function shape as the facilitation is not very strong and equally the masking even at the highest values is lower compared to other participant's suppression.

Participant D1 and D3 (Figures 7.3 & 7.5) tend to show the expected dipper shape with facilitatory handle and strong masking, especially participant D3. Participant D2 (Figure 7.4) somewhat shows the pattern of response as well.

The differences in pattern are also visible in the slope parameter data (see Figures 7.7, 7.8, 7.9, & 7.10). However, as detailed previously, some common patterns regarding relation between threshold values and slope values at baseline and low contrast level masks exist, which respectively tends to yield highest and lowest slope parameters.

7.2.2 Colour masks

The fitted threshold for each colour mask condition tested (see Table 7.2), following the method presented in Section 7.1.4, are presented in Figures 7.11, 7.12, 7.13 & 7.14. In each figure, I present the fitted threshold and the bootstrapped standard error of the threshold parameter (Section 7.1.4).The baseline, i.e. detection threshold without mask, is presented as a greyed area.

In the case of the chromatic masks, based on the results from Chen et al. (2000a) we would expect to have no effect at low chromatic contrast (if properly isoluminant) and suppressive masking at higher mask contrasts for colour are described as amplitude of cone contrast modulation (a) as defined in Chapters 4.

Participant D1 (the author) showed (Figure 7.11) no specific effect across all colour masks tested and this observer was tested with more colour contrast values than all other participants. This is interesting as it suggests a complete independence in the processing

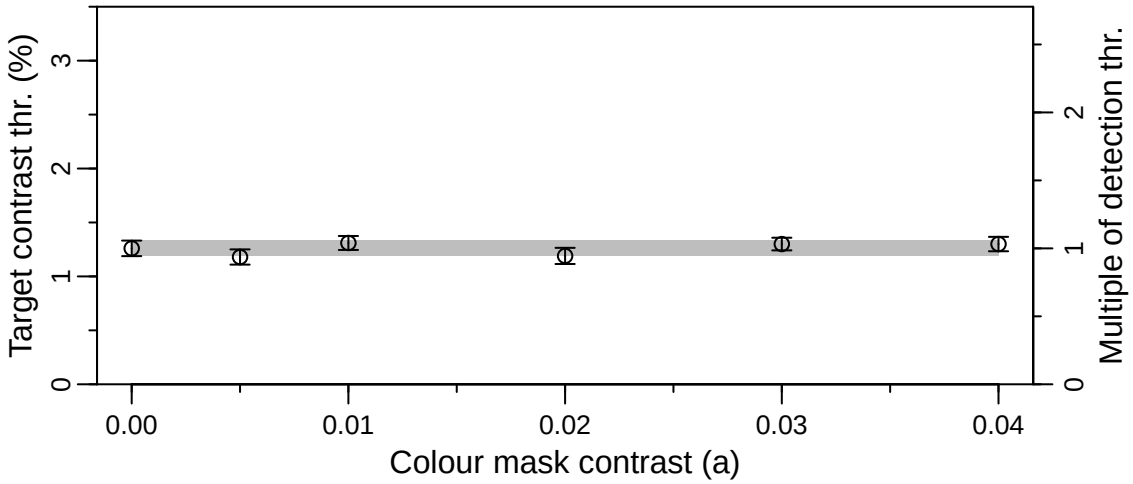


Figure 7.11: TvC response curve for colour mask participant D1/M1. Grey area: detection threshold estimates for luminance target without mask

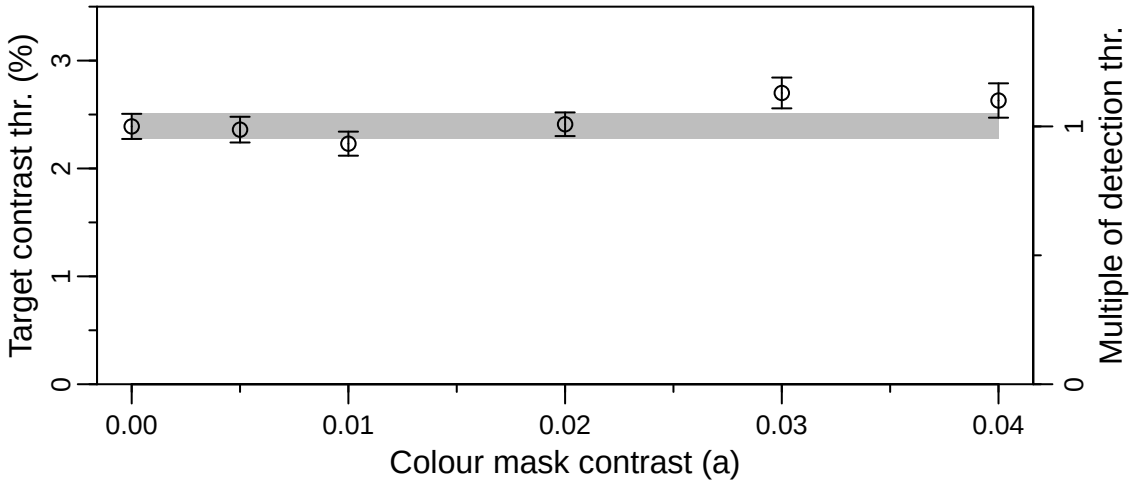


Figure 7.12: TvC response curve for colour mask participant D2/M6. Grey area: detection threshold estimate for luminance target without mask.

of luminance and orthogonal chromatic, at least at detection level.

From these results only participant D2 (Figure 7.12) shows the pattern expected from the literature. This participant does not show a strong pattern of suppression at the highest values, with the two highest colour contrasts tested delivering thresholds slightly above the baseline. However, for the luminance masks this participant (Figure 7.4) showed a similar strength of suppressive effect with high luminance contrast. The data from that participant was generally noisy.

Participant D3 (Figure 7.13) had two mask conditions showing suppressing masking effects, i.e. elevated detection threshold. This could be seen as matching the results for the luminance masks where most of the luminance masks conditions showed a suppressive effect, with only a facilitatory window from 5 to 7% (luminance mask contrast). However, no facilitation was found here, and the pattern of suppression does not grow with contrast.

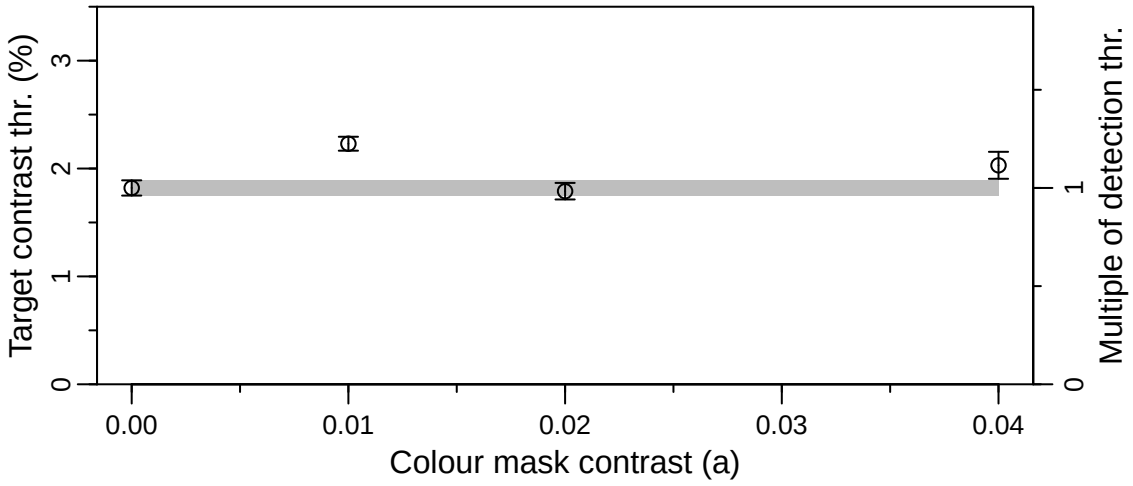


Figure 7.13: TvC response curve for colour mask participant D3/M3. Grey area: detection threshold estimate for luminance target without mask.

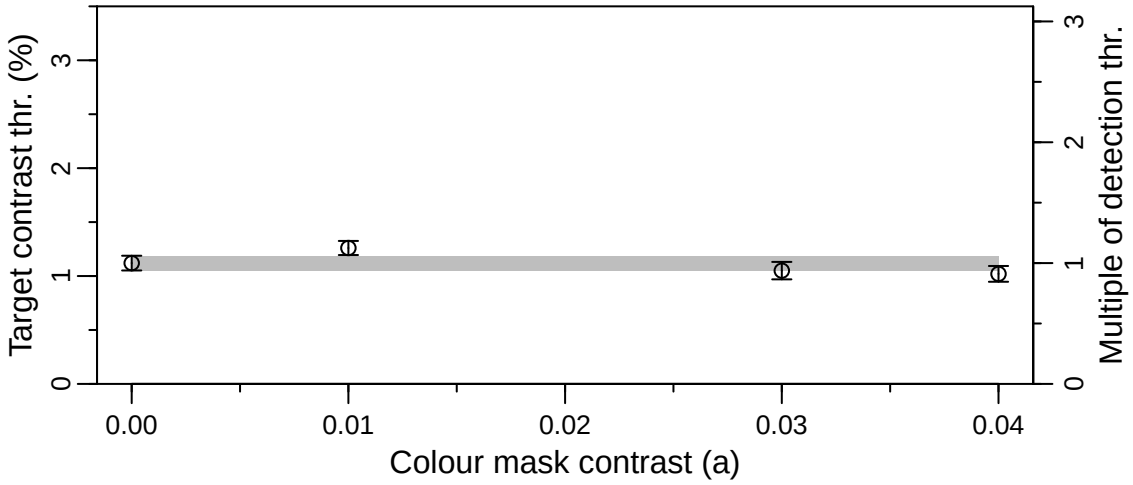


Figure 7.14: TvC response curve for colour mask participant D4. Grey area: detection threshold estimate for luminance target without mask.

Participant D4 (Figure 7.14) displayed only one mask condition with non-overlapping standard errors. It corresponds to the smallest colour contrast mask tested (for that participant) and is suppressive in nature. Compared to the luminance data (Figure 7.5), it seems to compare to the small contrast luminance effects.

Slope parameters

It can be seen, in Figure 7.15, 7.16, 7.17, & 7.18 that slope values tend to be homogeneous within participants. This is especially the case for participants D1 and D2 (Figures 7.15 & 7.16) and slightly less so for participants D3 and D4 (Figures 7.17 & 7.18). This homogeneity can be linked to the homogeneity of detection threshold result, and this further shows that the chromatic component seems no to be affecting detection of low contrast luminance gratings.

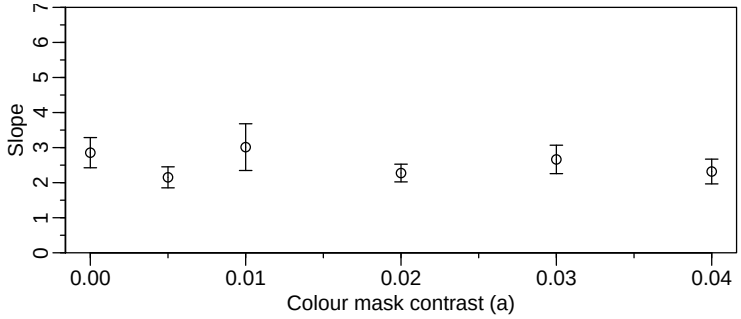


Figure 7.15: Slope parameter for colour mask, participant D1/M1.

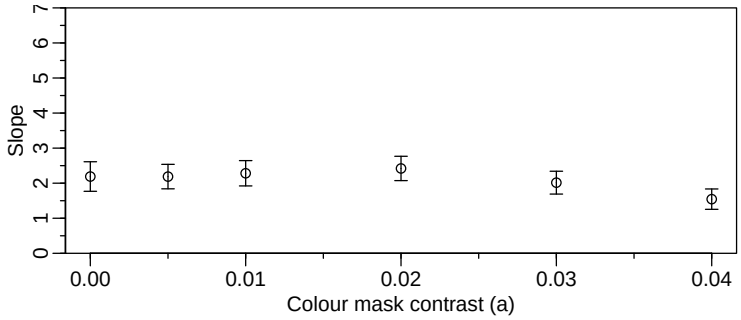


Figure 7.16: Slope parameter for colour mask, participant D2/M6.

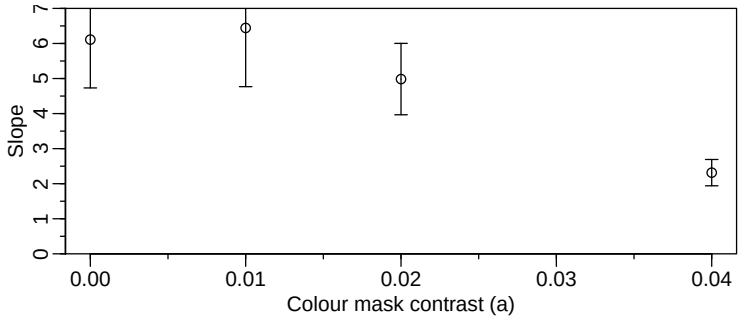


Figure 7.17: Slope parameter for colour mask, participant D3/M3.

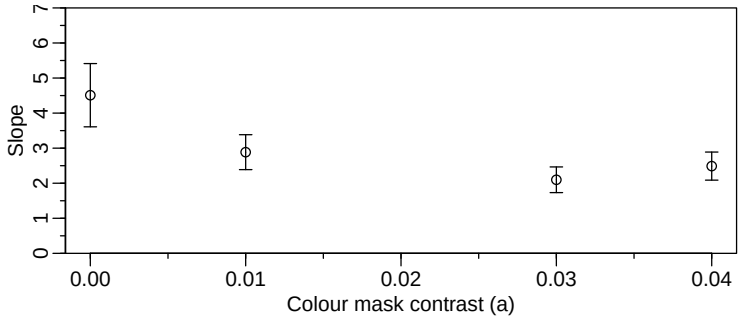


Figure 7.18: Slope parameter for colour mask, participant D4.

The mask contrast effect, in this section described as the typical dipper function effect, will be detailed in the next section using a computational model (family of models inspired by Foley, 1994). This will be useful to describe the shape of the TvC and get a better

understanding of the underlying mechanism.

7.2.3 Computational models

Introduction

In term of models, the dipper function behaviour is usually modelled using derivatives of the classic paper by Foley (1994).[6] This class of model has the form:

$$R = E^p/(I^q + Z) \tag{7.2}$$

R is the response of the detection mechanism; E is the excitatory drive, I is the inhibitory input; p, q, and z are parameters of the system that are assumed to be fixed. The commonly used parameters in the literature are $p = 2.4$ and $q = 2$ (Ross & Speed, 1991; Speed & Ross, 1992; Stromeyer & Klein, 1974). As I is the denominator, it is effectively having a divisive effect on the excitatory drive.

The parameters E and I responsible for excitation and inhibition are actually the contrast of the different elements in the stimulus. Therefore, we can effectively reformulate the above equation in terms of stimulus contrast; however, it is good to keep the terminology excitatory and inhibitory parameters as it will be helpful further on to understand more complex models. Foley (1994)'s model was inspired by *broadband divisive* effects found in the physiology and psychophysics (masking) literature.

Consequently, the excitatory input is more selective (i.e. tuned) and explains the psychophysical data that delivers a facilitatory dip tuned to the same spatial frequency, but suppressive effect similar for all frequency (broadband), see Chapter 3, Section 3.3.2 & Ross and Speed (1991). The compressive nature of the transducer function has also been modelled using a Naka-Rushton model (Ross & Speed, 1991, see Figure 3.10, p. 71).

According to Ross and Speed (1991), the masking can be understood by a change in the transducer function (or non-linear function), they also suggest that a shift of the transducer function towards higher contrast values is responsible for adaptation effects.[7]

One of the redefinitions of this model in terms of contrasts input, by Meese and Holmes (2007); Meese et al. (2007), is of following form:

$$R = c^{p_{ped}}/(z + c^{q_{ped}} + w * c^{q_{xom}}) \tag{7.3}$$

c_{ped} is the contrast of the pedestal, c_{xom} is the contrast of the cross-oriented mask. w is used as a weight on the effectiveness of the mask masking properties, it can vary when the mask and the target varies in orientation and/or SF (compare model by Ross & Speed, 1991, , equation 3, 4). p and q are the same parameters as in Equation 7.2.

[6] There is an earlier model by Legge and Foley (1980) with several derivatives in the literature, however it was found to be non-sufficient to explain more behaviours, by Foley (1994) and so is not presented here for conciseness. Prior models to Foley (1994)'s model also include work by Ross and Speed (1991), and Stromeyer and Klein (1974).

[7] According to Ross and Speed (1991), the mask might change the input contrast whereas the adaptor does not.

c_{ped} will be different in a target or no target interval. In the interval without target (R_m) c_{ped} represents the pedestal contrast only, in the interval with target (R_{m+t}), c_{ped} represents the contrast of the pedestal (Tped) plus the target (Tc). The threshold level of the system is the lowest value of target in which R_m can be distinguished (see Figure 3.9). Formally, R_{m+t} needs to be greater than R_m, however the sensitivity (k) of the system will affect how much more activity is needed in R_{m+t} to be detected:

$$R_{mask+target} - R_{mask} = k \tag{7.4}$$

In Foley's (1994) original model, we can see that the cross-oriented contrast is only present in the denominator so it will only have a divisive effect hence we cannot model cross-orientation facilitation within this framework. The model used in this chapter is based on the model by Meese and Holmes (2007) as described in the following section.

Fitting

The design for this experiment is mostly based on Chen et al. (2000a, 2000b) work however the model by Meese and Holmes (2007) was used to model cross-orientation effect as opposed to pedestal masking used in Chen et al. (2000a). Secondly, the fit was only performed for luminance-luminance condition as the results for the chromatic masks suggests no monotonic suppression effects (as expected from Chen et al., 2000a or dipper type functions.

The following equation was used taking into account the design of my experiment (i.e. no pedestal):

$$R = c_{target}^{p} * (1 + \alpha * c_{mask})/(1 + c_{target}^{q} + w * c_{mask}^{q}) \tag{7.5}$$

The parameter z, from Equation 7.3, was set to 1. As for Meese and Holmes (2007) and Meese et al. (2007) the modelling was performed using data normalized by detection threshold. Consequently, the parameter k (see Equation 7.4) was fixed at 0.5 (according to Meese & Holmes, 2007; Meese et al., 2007). This forces the fit to go through the baseline condition (Meese et al., 2007). The values of p and q were fixed (2.4 and 2 respectively), and the only two free parameters were α and w (as in Meese & Holmes, 2007).

The fits were obtained using numerical analysis to find the smallest values of c_{target} for which Equation 7.4 is true. The values obtained correspond to the threshold. The whole TvC function can be fitted by changing the masks values (c_{mask}). The sum of squares between the fit and the data was minimized using the function *fminsearch* from Matlab, this function implements the Nelder-Mead simplex direct search algorithm.

The data from participant D1 and D2 allowed for fits using this method (see solid lines in Figure 7.3 & 7.4). Participant D3 and D4 showed an anomaly in the second data point (first mask contrast after baseline) data, which require special arrangement. For these two participants, the data for the fit was not normalized by the detection threshold, instead

Table 7.3: Model Parameters

	Parameters					
Part.	p	q	a	w	z	k
D1/M1	2.4	2	0.428*	0.165*	1	0.5
D2/M6	2.4	2	0.044*	0.010*	1	0.5
D3/M3	2.166*	2.346*	0.114*	0.020*	1	0.5
D4	3.079*	1.491*	0.322*	0.206*	1	0.5

* free parameters

it was normalized by the second value of contrast.[8] This was done in order to make the model fit better. Therefore, the model is coerced to start at this value; furthermore, values of p and q were set as free parameters to allow for better fits. The best fit parameters are shown in Table 7.3.

7.3 Discussion

Chromatic effects

The principal research question was to search for chromatic (red-green) modulatory effects of masks on luminance targets. The evidence for such effects, in my experiment, is scarce. It is reasonable to conclude that in this configuration no cross-interaction between channels is happening, whether it is suppressive or enhancing.

However, some interesting information was gained from this experiment, especially if the results are compared to the ones on the luminance-matching task (see Chapter 6). Three of the participants tested previously performed this experiment (see participants M1(D1), M6(D2), and M3(D3) in Figure 6.4, p. 135). Amongst these participants M1(D1) showed a significant increase in perceived contrast at high (0.03) colour contrast, participant M3 showed no significant effect of colour (nominal decrease effect) and participant M6 showed a strong enhancing effect at high (0.03) colour masks. Most of these effects were tuned with luminance target contrast and disappeared completely at 2% contrast for participant D1/M1. The overall data from the suprathreshold matching experiment suggested that the effects were negligible at low levels of (luminance) contrast. However, this experiment was aimed at inspecting the possibility of suprathreshold mechanisms interfering at detection levels. The data observed do not find consistent interference.

A noteworthy observation is the absence of suppression, the suppression was previously observed by Chen et al. (2000a) in the cross-channel case. This suppression was present when both luminance and chromatic component were aligned (1 cpd). The suppression observed by Chen et al. (2000a), for their participant CCC, was of magnitude equivalent to a luminance of 4-8% contrast (Figure 1-top & Figure 8 Chen et al., 2000a), which already

[8] Note that an alternative method would have been to allow parameter k to vary.

yields significant suppression in luminance-luminance pedestal masking. It is possible that orientation difference (from iso-oriented for Chen et al., 2000a to orthogonal for my experiment) shifts facilitation or suppression to the right of the mask/pedestal contrast axis. This will be discussed in the luminance masking section.

The presence of luminance artifacts in this experiment is doubtful; however, the procedure for equiluminance was only performed once prior to the experiment. This is not the usual procedure for such experiment and equiluminant point should be re-evaluated on a regular basis. Chen et al. (2000a) did find some facilitation with nominal isoluminance, which disappeared once the equiluminant point was set. In my data, neither (consistent) facilitation or suppression was found, suggesting no artifacts.

Luminance effects

In iso-orientation (Chen et al., 2000a) as in cross-orientation (Meese & Holmes, 2007) the usual dipper function is found, however the facilitation dip is observed not at the detection threshold but later on, in both the above studies. In my data, this was found too, as visible in Figures 7.3, 7.5 & 7.6, where the detection threshold is represented vertically by a grey area. The facilitation dip happened beyond that level.

The amount of facilitation and suppression depends on the participant. (Meese & Holmes, 2007) showed that the facilitation was dependant on the temporal presentation as well as the SF. My stimulus temporal configuration matches more closely their 0.5 Hz stimulus, compare Figure 7.1 with their Figure 2 (Meese & Holmes, 2007, p. 129). Furthermore, the SF used in my experiment should lie between their 0.5 and 1 cpd results. At higher SF, suppression seems to be rarer and facilitation seems to be prevalent albeit shifted to higher contrast values (Meese & Holmes, 2007).

Two participants (D2 & D3) demonstrated an anomaly in which detection levels, when the mask was at or below threshold, were higher than expected. This prompted a specific fitting model (see Figure 7.5 & 7.6), interestingly this is somewhat visible for participant T.S.M (Figure 5-b Meese & Holmes, 2007, p. 131). The difference between my experiment and Meese and Holmes (2007) is that I sampled the lower contrast values to a larger extent. It is possible that in that range other mechanisms are present.

The general conclusion is that my data corroborate the presence of facilitation in orthogonal luminance masks as well as the more common suppression effect.

Fitting results

Only the luminance mask was fitted and so the mechanism describe in this section only concerns within-luminance channel interactions.

Firstly, I will detail the fit for participant D1. Using p and q parameters (Equation 7.5) commonly used in the literature the fit on parameters α and w matched what would be expected from the literature (Meese & Holmes, 2007). A strong α value corresponds to a strong input from the orthogonal mask to the target and a strong w value correspond

to a strong normalization process.[9]

Participant D2 does not show a strong effect of the mask, consequently the two parameters (see Table 7.3) are low, allowing for a shallow facilitation followed by a weak suppression at higher mask values.

Participant D3 and D4 required adjustment of the normalization, although as noted earlier, setting k as free parameter is a possible solution too. The fit for participant D4 involves a strong p and lower q (see Table 7.3). α is larger than w as is the case for all the other fits. Finally, participant D3 best fit had unexpected results of having $q > p$ which is not usually found in these kind of models, but the mask parameters coerce the fit into the usual dipper shape.

The fits demonstrated that the strength of orthogonal effects in term of inhibition and excitation could vary between participants while still preserving the generic result that is the dipper function.

Expansion of the model would include using less free parameters, such as relaxing k and z. The effect of surround versus centre variation within the model affecting those parameters has been studied by Meese et al. (2007). It would be an interesting addition to the experiment to add surround modulations to the stimulus.

Conclusion

The luminance results are in line with expectations. Extension of the experimental protocol could look into orientation tuning and its effect on the parameters of the current model in order to better understand the tuning of these orthogonal effects (e.g. Ross & Speed, 1991). This would be especially interesting for the chromatic case, as a measure of the suppression orientation tuning would bridge my data and the one by Chen et al. (2000a).

Furthermore, once again my data demonstrates variation between participants; it would be interesting to characterise the extent of these variations amongst a larger population however the task might be difficult due to the number of trials required. A shortening of the number of conditions (masks) should be considered.

There were no effects of a chromatic mask in these experiments. In order to find chromatic effect it is possible that the mask needs a luminance edge to interact (see intro Psychophysics), consequently the use of mask composed of pedestal luminance plus variable chromatic masks should bring valuable information about the necessity of co-alignment for interaction.

The spatial extent of the facilitatory and suppressive effects as well as the relation of presentation type (e.g. binocular, dichoptic or monocular) should also be studied as these might yield information on the physiological underpinning of the cross-channel interactions.

[9]The tuning of these, in term of orientation, has been discussed in the Psychophysics introduction, Chapter 3.

Chapter 8

Isoluminance analysis[1]

This section is complementary to the literature review (Chapter 2) on physiology but also on units and light description (Chapter 4). It describes the isoluminance experiments used in this thesis and as such is an extension on Methodology, and a data chapter as well. The aim of the study was not to study isoluminance per se but an interesting body of research has emerged from my experiments on the subject and I believe this provides an interesting tangent into the study of luminance and chromatic channels.

This chapter introduces the reason for the need to establish equiluminance and then introduces the methodology of equiluminance calibrations. I then report the results of equiluminance experiments for this thesis and discuss the implications for luminance and colour perception, but also for further studies of equiluminance.

8.1 Introduction

8.1.1 Physiology

Generally, the contribution of L and M inputs to the luminance signal are not equal (Gunther & Dobkins, 2002). This can be due to two main factors, firstly the numbers of L and M cones available, which varies between observers (Kremers et al., 2000), as well as between primate species (Dobkins, Thiele, & Albright, 2000). The second factor is the post-receptoral weighting that computes the luminance channel. We have seen in the physiology review, Chapter 2, that the ganglion cell receptive field is determined by both connectivity preferences and the availability of cones (which depends on the cone ratio); this was shown by Field et al. (2010). Furthermore, Brainard et al. (2000) showed that the L and M cone ratio was preserved in the ERG (Electroretinogram) signal; the ERG signal is a sum of all the ganglion cell activities (see Physiology). By looking at the amplitude of the signal over a range of wavelengths Brainard et al. (2000) were able to find the equivalent ratio of L to M cone producing this sensitivity curve. Finally, they compared it with the ratio obtained from psychophysics and found good agreement.

[1]Parts of this chapter have been published, the publication, (Clery et al., 2013), is attached in Appendix B.

8.1.2 Isoluminance and Isochromaticity

Isoluminance, also sometimes referred as equiluminance, means equal amounts of light. It is a useful concept when trying to isolate luminance and colour channels, or in the case of this thesis, when we see how the two interact further along the processing stream.

This concept assumes that luminance and colour are processed by separate systems (at least in the early stages), see physiology review (e.g. Livingstone & Hubel, 1988). If one wants to explore only the luminance channel, one needs to change the intensity of the light (in cd/m^{-2}) without changing the relative colour distribution (or light spectrum). This is called an isochromatic, or achromatic, stimulus.

The term achromatic might be misleading, as it suggests a lack of colour; what is actually meant is a lack of chromatic variation within the stimulus. However when I refer to the achromatic (or luminance channel) here it is meant that this channel does not encode colour information. Consequently, an achromatic stimulus does modulate the activity of the channel but a *purely* chromatic stimulus (i.e. an isoluminant stimulus) would not.

8.2 General methodology

Now as we have seen in the Methods, Section 4.3.1, in order to create a chromatic stimulus we create a stimulus that modulates L and M cones out of phase (and keeps S cones at a constant level of activity). A first attempt at isoluminance can be made by producing a nominally isoluminant stimulus.

8.2.1 Nominally isoluminant stimuli

This can be done by moving back from the cone contrast description (that I used preferably throughout this thesis), to the experimental setup colour space. In this colour space, I am mostly dealing with a red and green gun, or red and green light. It is nominally possible to create an isoluminant stimulus by modulating the red and green guns (or lights) so that their light intensity (in cd/m^{-2}) sums to the same overall value and is constant across space on the stimulus.

This method is based on the photopic standard observer function (see Methods, photopic functions, Figure 4.2, p. 78), hence the measure in cd/m^{-2}. It has been recognized that using solely this method, or an analogous method that does not result in an observer defined isoluminant pattern, will result in luminance anomaly in the stimulus (see Dobkins, Anderson, & Kelly, 2001, for example). By using the neurophysiology technique of EEG, it is possible to detect such anomalies as the luminance signal is usually triggering a fast positive signal and chromatic signals are associated with slower negative signals in the EEG (see for example Parry & Robson, 2012, for isoluminant, yellow-blue stimuli and EEG signature of luminance artefacts).

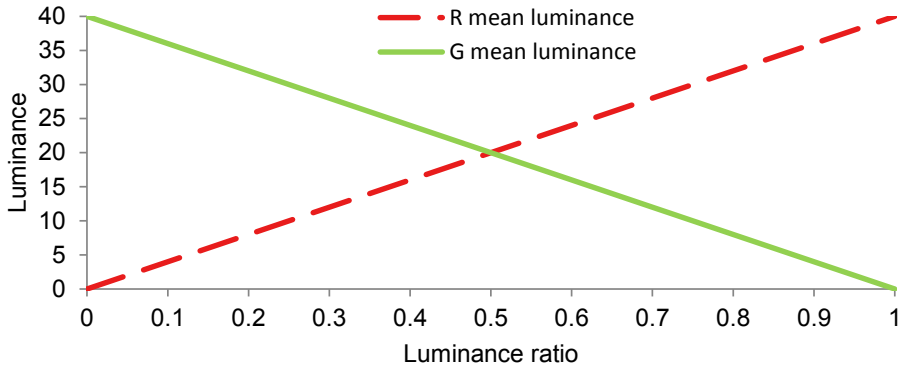


Figure 8.1: Luminance ratio and gun's chromaticity values. This figure illustrates the relationship between the gun mean luminance and the ratio obtained from Equation 8.1. Nominal isoluminance is present when the two lines intersect, however perceived isoluminance might vary around this value.

The differences in cone numbers has be linked to stochastic phenomenon in gene expression, with possible epigenetic silencing involved (for a review see Neitz & Neitz, 2011). This differences in cone ratios in the retina as well as possible differences in cone sampling from the ganglion cells both result in isoluminance being different from one observer to another. This is creating the need for individual calibration that varies around the nominal isoluminant stimuli.

8.2.2 Perceptually isoluminant stimuli

A simple methodology to counteract this problem consists of adjusting the red and green components of the stimulus until they become perceptually equivalent in luminance. This can be done using the setup colour space but equivalently we can adjust the depth of chromatic contrast of L and M cone if one defines the stimuli in cone space or cone contrast space.

In order to define the stimulus directly in gun luminance intensity it is common to describe the red-green ratio in the following manner:

$$L_R/(L_R + L_G) \tag{8.1}$$

In this equation, L_R is average luminance of the red gun value and, L_G is the equivalent for the green gun. The relationship between the ratio and the two guns is linear as can be seen in Figure 8.1. The sum $L_R + L_G$ is fixed at a specific value, so that when adjusting the relative values an observer will add as much luminance in the one gun as is taken from the other one (see Figure 8.1). To the best of my knowledge the first study to use such a paradigm was devised by Mullen (1985), in her seminal work on colour channels.

For reference, following this definition above, a nominally isoluminant stimulus yields a ratio value of 0.5, it is the middle point in Figure 8.1. Using an isoluminance probing technic (such as minimum motion, minimum flicker and minimum border) that varies the

ratio in Equation 8.1 it is possible to determine the amount of red and green luminance (and gun values) for a particular individual.

This is a practical way to produce individually defined isoluminance. Using this method the contrast of the stimulus is defined using a definition of contrast such as Michelson or Weber (see Methodology Chapter 4). This is done as a function of red and green phosphor modulation.

Consequently, the definition of the stimulus consists of two luminance contrasts values, one for the green and one for the red. For reference again, the nominal isoluminance is when the two gratings add up to the same value of luminance across the whole space. The luminance contrast amplitude of the two stimuli (red and green light) are equal but their mean nominal values can be adjusted relatively (See Figure 3 from Mullen (1985) for detailed explanation). At ratio of 0 or 1 the stimulus a pure green or pure red modulated sinusoid, see Figure 8.1.

The definition of these stimuli is based on red and green phosphors; an alternative method is to express the stimuli in term of cone modulation (both methods are accepted Brainard, 1996; Vimal, 1997). In Mullen and Losada (1994), a similar method was used, the stimulus luminance profile (L) was given by the following equation (p. 3 in original paper) describing both red and green phosphors:

$$L(x) = M_r + M_g + M_r(C + \Delta C)\sin fx \pm M_g(C - \Delta C)\sin fx \tag{8.2}$$

M_g & M_r are the average luminance across the pattern, x the horizontal position on the stimulus[2], f the spatial frequency (SF), C is the contrast amplitude, ΔC is used to add an additional luminance component to the stimulus.

This definition was designed to create luminance stimulus (additive mode between g & r, and $\Delta C = 0$) or chromatic stimulus (subtractive mode between g & r, and $\Delta C = 0$). The isoluminance was obtained, as defined above in this section, by adjusting the relative M_g & M_r (but $M_g + M_r$ is constant over all conditions). C is equal between red and green phosphors. This equation also allows creating overlayed luminance and achromatic stimuli by having $\Delta C \neq 0$. Note that the equation to create stimuli in cone contrast is inspired partially from this design; instead of using the mean luminance of red and green (which is done using standard observer luminosity function), I used an analogue of ΔC (b) to modulate cone contrast. The methodology for creation of isoluminant stimuli in cone-contrast space, which was used in my experiments, is defined in the following section.

8.2.3 Perceptual isoluminance in cone-contrast space

In this section, I will apply the methods described above to the cone-contrast colour-space I used for my experiments. This explains in more details the math and methodology behind the adjustment required by participants in the specific cone space defined in the

[2]This produces a vertical sinusoids, the equation can be extended to allow other orientation, see General Method section Equations 4.7, 4.8 & 4.9.

Methodology Chapter 4. First, I will give a brief summary of how our stimuli are defined in term of contrast; the full details on stimulus generation are available in the Methodology chapter. Then I will detail the procedure required to make them equiluminant using the mathematical definitions of our space, which introduces variation of the original definitions.

Summary of cone contrast space The stimuli are defined in term of cone-contrast modulation. The cone contrast is a measure of variation from the mean background intensity, this definition is based on Weber contrast (this is also true for Michelson contrast). Consequently, using this cone contrast metrics, sinusoid and square wave patterns of similar highest and lowest values have the same contrast.

The variation of cone contrast used in the creation of periodical stimuli for the three types of cones are the values A_L, A_M and A_S, A signifying amplitude modulation of cone contrast. In order to create red-green modulated stimuli I modulate L and M cone contrast out of phase, the values are transformed into cone excitation (see Methods chapter 4) and then the excitation values are transformed into gun values to be displayed on the screen[3]. As a reminder the Equation for L, M & S cone contrast modulation (from the method section) are reproduced here:

$$L_{Csin}(x,y) = (\pm A_L)\sin([(ycos\theta) - (x\sin\theta)] * 2\pi f) \tag{8.3}$$

$$M_{Csin}(x,y) = (\pm A_M)\sin([(ycos\theta) - (x\sin\theta)] * 2\pi f) \tag{8.4}$$

$$S_{Csin}(x,y) = (\pm A_S)\sin([(ycos\theta) - (x\sin\theta)] * 2\pi f) \tag{8.5}$$

However, as we have seen earlier in this section, L and M contributions to the luminance channel might be unequal. Consequently, when L and M cone contrast are out of phase the stimulus might not be equiluminant. If we were to represent the responses of the red-green and luminance channel on top of the stimuli, we would see a strong sinusoidal response from the chromatic channel (red-green) and a small sinusoidal response from the luminance channel either in-phase with the red or green bands of the stimuli depending of which are perceived as brighter. A perfect isoluminant stimulus would have a sinusoidal response from the chromatic channel of interest and flat response from the luminance channel[4].

L and M Cone adjustment In order to correct for this and create equiluminant stimuli, I require the participants to adjust the respective amplitude values of L and M cone contrast. This adjustment corresponds to the varying of the value b in Equations 8.6 & 8.7 below. By doing this, an equal amount of cone contrast is effectively subtracted from

[3]The gun values then undergo gamma correction; this is the last step providing the gun values to be displayed on the screen.

[4]The concept of psychophysical channels being silenced should not be interpreted as a physiological channel being completely silenced, it would correspond at the lowest response from the population; for a more detail discussion see Regan (1991).

one cone type and added to the other type. Thus, the sum of the two contrasts stays constant. This is the equivalent of the ΔC parameter presented in the previous section on isoluminant stimuli (in device based stimuli definition).

Essentially this equation adds an amount (b) of L contrast and subtracts the same amount (b) of M contrast or the opposite; and the parameter a is the nominal equiluminant contrast value (see section 8.2.1). However, as M cone contrast is given a negative value (i.e. to create a phase offset producing red-green modulation) the equations take the following form:

$$A_L = a + b \tag{8.6}$$

$$A_M = -a + b \tag{8.7}$$

$$A_S = 0 \tag{8.8}$$

The resulting A_L, A_M and A_S are then plugged in the equation defining our sinusoids see Equations 8.3, 8.4 & 8.5. This describes the cone contrast modulation of each cone class. It is then transferred into nominal cone excitation values (using the cone contrast definition), which in turn is transformed into gun values (using method outlined by Brainard et al., 2002). Finally, the gun values are gamma corrected. The method is described in more detail in the main Method section.

Note that there are alternative ways to produce out of phase cone excitation. One way is to add a phase shift parameter in the sinusoid Equation. In this case, the amplitudes can both be set in positive values and the adjustment would be defined as:

$$A_L = a + b \tag{8.9}$$

$$A_M = a - b \tag{8.10}$$

$$A_S = 0 \tag{8.11}$$

Determining modulatory parameter b When isoluminance is not required (achromatic stimuli), A is equivalent to a and L, M and S are modulated by the same amount and in-phase. Consequently the parameter a was used to describe contrast in my experiments for both luminance and chromatic, but for each value of a (i.e. cone contrast amplitude before modulation) we need to determine a value of adjustment b. The mathematical relationship between a and b is an empirical question. The aim of the equiluminant experiments that are presented in the rest of this chapter was to determine for every contrast amplitude (a) that I wanted to display, how much modulation from nominally isoluminance was necessary (b) so that the stimuli would be equiluminant for a specific participant.

Describing stimuli using specific L/M pairs values is also common in the literature (see Wuerger & Landy, 1993), however in order to simplify the read of the previous

experimental Chapters it seemed more appropriate to give the full description of the L/M excitation pairs separately.

Worked example Figure 8.2 shows a worked example. In this example, we measure the relationship between A_L and A_M values that deliver perceptual isoluminance for two different behaviours. In one case nominally isoluminant (dotted line) and one case in which an observer requires 4 times more contrast for the M cones than for the L cones to make a stimuli isoluminant (data points + solid line).

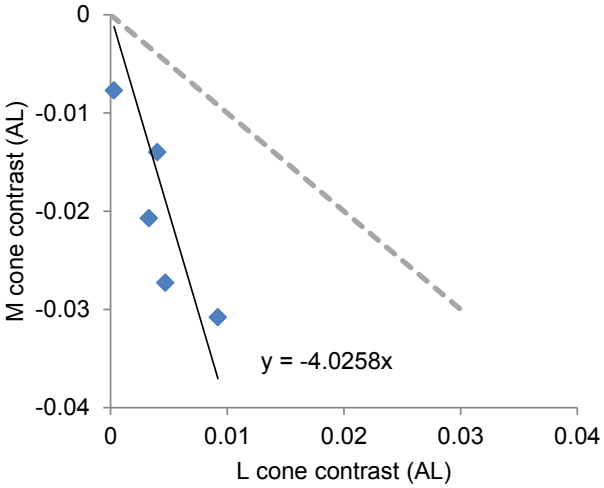


Figure 8.2: Relationship between A_L and A_M (L and M cone contrast amplitude modulations).Dotted line: line representing a (virtual) observer with equal L and M cone contribution to luminance, i.e. equal cone contrasts does not generate any luminance percept. Solid line: line representing a participants requiring 4 times more M contrast than L contrast (when out of phase) to be isoluminant. Diamonds represents actual data from a participant showing that behaviour, the full dataset of results are presented in a following section.

A complete description of the stimuli would involve giving the nominal A_L and A_M values. However if one knows the values of a and b, one can find the individual cone modulation easily using Equations 8.6 & 8.7, as was done to create the stimuli for the experiments (and in the Equiluminance setting experiment).

This is useful for contrast computation; in our experiments we always refer to contrast as parameter a, independently of the isoluminance adjustment, and as for luminance there is no adjustment, it corresponds directly to A. We use the following section to describe the root mean square cone contrast for our colour stimuli.

In the following description of the Equiluminant methodology and results, the parameter definition follows Equation 8.9 & 8.10, however the conclusions we draw apply to the alternate method of generating chromatic stimuli.

8.2.4 R.M.S cone contrast

In the sets of experiments presented in this thesis the contrast of each colour stimulus is described using a, as it is observer independent. However, there are other alternatives, such as specifying all L contrast M contrast pairs individually (e.g. Wuerger & Landy, 1993), or using a summary statistic such as R.M.S cone contrast, with L and M cone only (e.g. Gunther & Dobkins, 2002) or using L, M, and S (e.g. Giulianini & Eskew, 1998).

Giulianini and Eskew (1998) defined contrast as the Euclidean distances (here modified to match the notations used previously in this Chapter and Equation 4.6):

$$R.M.S = \sqrt{\frac{[\Delta\epsilon_L/\epsilon_{bL}]^2 + [\Delta\epsilon_M/\epsilon_{bM}]^2 + [\Delta\epsilon_S/\epsilon_{bS}]^2}{2}} \tag{8.12}$$

$\Delta\epsilon_L$ is the cone excitation difference from baseline for cone type L, ϵ_{bL} is the cone excitation background value. The same nomenclature is used for M and S cones. $\frac{\Delta\epsilon_L}{\epsilon_{bL}}$ as defined in the General Method chapter 4 (Equation 4.6) is the Weber cone contrast (cone contrast as defined by Cole & Hine, 1992). Remember that these values can be positive or negative. These values of contrast (L_c, M_c and S_c) were modulated differently (according to space and time) depending on the experiment.

As I used recurrent stimuli (sinusoids and square-wave) the definition of the stimulus included a sine-wave and an amplitude modulation for each cone type (A_L, A_M, A_S). This amplitude modulation therefore is between -A to +A. This directly translates to a cone contrast (Cole & Hine, 1992) from -A to +A. In the case of square-wave pattern each stripe types (e.g. red and green) the cone contrast will be identical but of opposite sign. Therefore the contrast is constant over space, with sinusoids the contrast oscillates between equal values of A (of opposite signs).

Consequently, I can reformulate Equation 8.12 as:

$$R.M.S = \sqrt{\frac{[A_L]^2 + [A_M]^2 + [A_S]^2}{2}} \tag{8.13}$$

In addition, with our red-green chromatic stimuli by definition $A_S = 0$, thus Equation 8.12 reduces to:

$$R.M.S = \sqrt{\frac{A_L^2 + A_M^2}{2}} \tag{8.14}$$

Gunther and Dobkins (2002) defined (similar to Reisbeck & Gegenfurtner, 1998) R.M.S contrast as (using adjusted notations, with highest cone excitation C_{max} and lowest C_{min}):

$$R.M.S = \sqrt{\frac{[L_{max} - L_{min}/L_{max} + L_{min}]^2 + [M_{max} - M_{min}/M_{max} + M_{min}]^2}{2}} \tag{8.15}$$

I have demonstrated previously (General Methods, Chapter 4) that $[L_{max}-L_{min}/L_{max}+L_{min}]$, which is the definition for Michelson cone contrast, is equivalent to A_L (and the same goes for A_M). As $[L_{max} - L_{min}/L_{max} + L_{min}]^2$ is equivalent to A_L^2 and $[M_{max} - M_{min}/M_{max} + M_{min}]^2$ to A_M^2, I can then reformulate the Equation 8.14 as follows:

$$R.M.S = \sqrt{\frac{A_L^2 + A_M^2}{2}} \tag{8.16}$$

We can appreciate that in this condition this yields a definition identical to Giulianini and Eskew (1998) defined above (Equation 8.14) in the case of a stimulus with no S-cone modulation.

In the previous section I described how to obtain individual cone contrasts (A_L & A_M) from modulation b, $A_L = a + b$ and $A_M = -a + b$. I can then reformulate Equation 8.16 as follows:

$$\sqrt{\frac{A_L^2 + A_M^2}{2}} \tag{8.17}$$

$$\sqrt{\frac{(a+b)^2 + (a-b)^2}{2}} \tag{8.18}$$

$$\sqrt{\frac{a^2 + 2ab + b^2 + a^2 - 2ab + b^2}{2}} \tag{8.19}$$

$$\sqrt{a^2 + b^2} \tag{8.20}$$

This is the root mean square formulation at the level of cone contrast amplitude modulation. Using this definition for a fixed a, any modulation b will have an effect on the cone contrast; however we expect it to have a minor effect. It is possible to replace the values of contrast in our results section with the R.M.S. definition (or any other[5]), and I have shown here how to do so.

However the transition between RMS and using a should not be critical as the relationship between our definition of contrast and the R.M.S are both monotonically increasing. As a gets larger so does the R.M.S values and we expect b to change monotonically with increasing values of a without becoming greater than a. This expected monotonicity has never been tested in the literature formally. To the best of my knowledge, in the literature experimenters tend to test isoluminance at large contrast values and then extrapolate for smaller values. This is to be expected because larger chromatic modulation, if not properly isoluminant, will create larger luminance artefacts, and when we approach smaller values we expect smaller artefacts. At 0 colour contrast modulation, there is no need for luminance correction so we expect the intercept to pass through zero.

Concerning the relationship between adjustment (b) and contrast before adjustment (a), I tested the adjustment levels for every contrast level required in latter experiments (and as we will see in the results section, the expected linear relationship was found).

[5]For example McKeefry, Parry, and Murray (2003); Medina and Diaz (2006) $\sqrt{[A_L^2 + A_M^2 + A_S^2]/3}$

Providing the values of a and b one can transform the contrast values into one's metric of choice if necessary; however the results and conclusions are not dependent on the contrast metric used. Furthermore, it is convenient to use this metric (i.e. parameter a) as when L, M, & S cone contrasts (Lc, Mc and Sc) are identical (b=0), this metric is equivalent to the Michelson contrast (see proof in Method section Section 4.3.4, p. 89).

8.2.5 Additional comments regarding contrast depth

The method chosen to make stimuli equiluminant as defined by adjustment of parameter b in Equation 8.6 & 8.7 (somewhat similar to the method used in Mullen (1985) but with cone contrast) has the advantage of preserving the sum of the cone contrast modulation. Following these equations, a negative b would make the L contrast smaller and relatively the M contrast larger, and vice versa with positive values. This is a symmetrical adjustment of contrast on those two types of cones. Using this methodology preserves the overall contrast depth, which is if we consider the overall contrast depth to be the summation of contrasts between L and M cones.

To put it formally, in our methodology $|A_L| + |A_M|$ is constant, furthermore $|A_L| + |A_M| = 2a$. a is the value of contrast amplitude that I report as contrast both for luminance and chromatic contrast. In the Methods Section 4.3.4, I have shown that the amplitude value A_C is equivalent to the Michelson contrast for cone type C. By extension, I have shown that it corresponds to the Michelson contrast for luminance stimuli. This produces an achromatic modulation $A_L = A_M = A_S = a$ and $b = 0$. As a consequence I always report a as the contrast value for both types of stimuli, but as I have demonstrated in the previous section any other contrast metric can be derived knowing a and b.

Here, I work out an example showing the constancy $|A_L| + |A_M| = 2a$. For example, if $b = 0$ and $a = 0.05$ on a red stripe, $A_L = +0.05$ and $A_M = -0.05$; and on a green stripe, $A_L = -0.05$ and $A_M = +0.05$. The maximum, $|A_L| + |A_M|$, is 0.1. For $b = 0.01$ and $a = 0.05$ on a red stripe $A_L = +0.04$ and $A_M = -0.06$ (green stripe: $A_L = -0.04, A_M = +0.06$). The maximum, $|A_L| + |A_M|$, is preserved at 0.1.

This is different from an alternative procedure to obtain isoluminance in which the L contrast is fixed and the M contrast is varied, or in which case the red gun is fixed and the green gun is modulated, e.g. Johnson et al. (2001). Furthermore, it is of importance as contrast is changing the saliency of the stripes (Blaser, Sperling, & Lu, 1997), and this thereafter might be used to compute isoluminant motion (Z.-L. Lu, Lesmes, & Sperling, 1999). I will discuss motion and isoluminance in a latter section.

The fixed contrast I just described might however be dependent on the type of contrast metric used (see Chapter on methodology). Using a squared definition of contrast, such as R.M.S, would suggest variation of chromatic contrast with b. Such variations are minimized using our method. The implications of individual isoluminance points and therefore individual differences in cone inputs to the colour and luminance channels are difficult to take into consideration in term of using an accurate measure of contrast magnitude for

every participant. Using the methodology I outlined in the Methods chapter and in this chapter, I tried to minimize the impact of this difficulty.

8.2.6 L:M cone Ratio metrics

This section delves into the cone ratio calculation; this is not part of the Methodology on isoluminant determination per se. This section details how we can estimate the cone ratios using the metrics defined before and consequently is an extension to the physiology of cone ratio and stimuli methodology.

Mullen (1985) used luminance ratio to define the relative contribution of red and green stripes to the luminance channel. It is possible to devise a metric performing a similar computation but for cone excitation. Additionally one can also calculate L:M cone ratio from contrast amplitude values. Dobkins et al. (2000); Gunther and Dobkins (2002) gave a definition that uses L and M cone nominal excitation level (as opposed to contrast values) at their maximum and minimum values, which corresponds to red and green stripes (depending on the cone type). In their paper they start with the premise that at equiluminance the sums of inputs from L and M cones to the green stripes is equivalent to the sum of inputs of L and M cones in the red stripes towards the luminance channel, mathematically as follow (reformulated using notations previously used):

$$w_L L_{max} + w_M M_{min} = w_L L_{min} + w_M M_{max} \tag{8.21}$$

w_L is the weighting of L input to the luminance channel and w_M the inputs from the M cones. The two stripes are represented by different sides of the equation. They define the cone ratio as the ratio of cone weighting, w_L/w_M. I provide below the proof of the derivation of their equation for w_L/w_M or L:M ratio calculation (not presented in the original paper).

$$\begin{aligned} w_L L_{max} + w_M M_{min} &= w_L L_{min} + w_M M_{max} \\ w_L(L_{max} - L_{min}) &= w_M(M_{max} - M_{min}) \\ w_L/w_M &= ((M_{max} - M_{min})/(L_{max} - L_{min})) \end{aligned} \tag{8.22}$$

Hence to obtain the L:M ratio I use (as described in Dobkins et al. 2000; Gunther and Dobkins 2002):

$$(M_{max} - M_{min})/(L_{max} - L_{min}) \quad = \quad \text{L:M ratio} \tag{8.23}$$

Recall from the main method section, in the demonstration of Section 4.3.4 that $M_{max} - M_{min}$ corresponds to the full amplitude of the modulation hence is equivalent

to $2A_M$ times the background excitation level ϵ_{bM} (and $L_{max} - L_{min} = 2A_L * bc_L$). Hence, I can rewrite the Equation 8.23 as:

$$\begin{aligned} \frac{2A_M \times \epsilon_{bM}}{2A_L \times \epsilon_{bL}} &= \text{L:M ratio} \\ \frac{A_M \times \epsilon_{bM}}{A_L \times \epsilon_{bL}} &= \text{L:M ratio} \\ \frac{A_M}{A_L} \times \frac{\epsilon_{bM}}{\epsilon_{bL}} &= \text{L:M ratio} \end{aligned} \tag{8.24}$$

As we can see, the equation can be reorganized as the ratio of L/M cone contrast times the L/M background excitation ratios. In Figure 8.1, in the previous section, I described the cone contrast ratio and it is relatively easy to calculate once one knows the parameters a *&* b from Equation 8.6, the ratio of contrast is actually $(a-b)/(a+b)$. As it stands this ratio is independent of the background values.

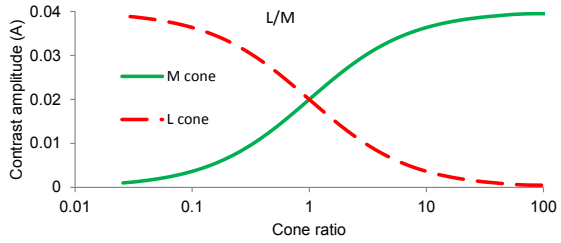
However, to comply with Gunther and Dobkins (2002)'s definition I have to multiply it by the ratio of cone excitation at background level. In our experiment bc_M and bc_L are quite close, as I am using a grey background, the relative ratio is $bc_M/bc_L = 0.906$. This ratio is constant across all our conditions and the same between experiment I (part 1 & 2) and experiment II and III (Chapters 5, 6, & 7) as I used the same values of background.

An illustration of the behaviour of the cone ratio equations can be found in Figure 8.3. The alternative metrics would be to compute the cone ratio using a formula akin to the one describe in Equation 8.1, $L/L+M$, from Mullen (1985). This can be calculated as $1/[(A_L/A_M)+1]$. The relationship between a x/y ratio and a $x/(x+y)$ is described as follows:

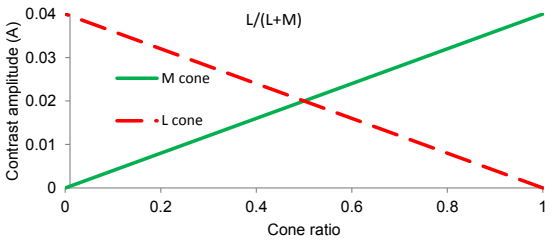
$$\frac{x}{x+y} = \frac{x/x}{(x+y)/x} = \frac{1}{1+(y/x)} \tag{8.25}$$

It is therefore possible to calculate this other ratio using the original ratio as describe in Equation 8.24 (derived from Gunther & Dobkins, 2002 and Dobkins et al., 2000). Having close nominal excitation levels for L and M for the background mitigates the shift on the cone ratio. The definition of cone ratio purely based on cone contrast is however invariant of background value.

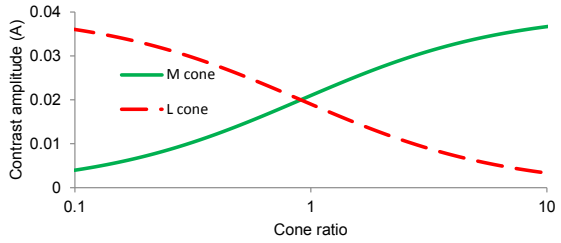
To summarize, it is possible to characterize the input ratio of the L and M cone to the luminance channel. This information is a by-product of constructing an isoluminant stimulus. This ratio comes from early literature (Mullen, 1985), that gave luminance ratio of red and green stimulus to the luminance channel. However, this was improved upon by the use of direct cone defined stimulus (Dobkins et al., 2000; Gunther & Dobkins, 2002) that allowed estimated of L and M cone inputs. To conclude this section I will note again the interesting results from Brainard et al. (2000), showing that the L/M cone ratio has an effect on the highest spectral sensitivity of the ERG and the two were linked to the



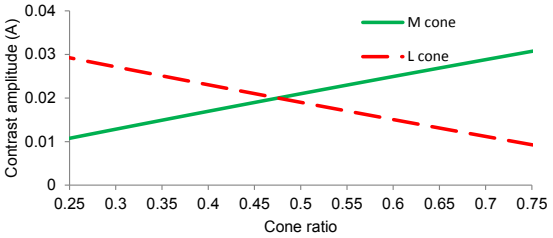
(a) Relationship of L and M contrast amplitude as a function of cone ratio, A_M/A_L, the first half of the Equation 8.23. Equal L and M cone contrast gives a ratio of 1. Higher cone contribution in the favour of M cones leads to lower cone ratio and the opposite for L cones. The cone ratio is presented on a logarithmic scale.



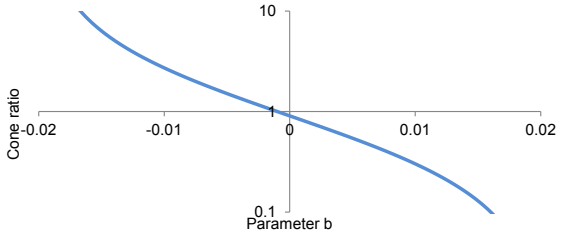
(b) Relationship of L and M contrast amplitude as a function of normalized cone ratio (L/L+M), $A_M/(A_L + A_M)$. Equal L and M cone contrast gives a ratio of 0.5. Higher cone contribution in the favour of M cones leads to lower cone ratio and the opposite for L cones. The cone ratio is presented on a linear scale.



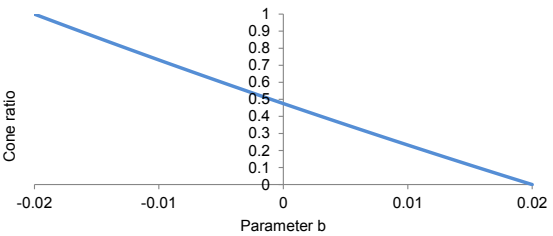
(c) Relationship of L and M contrast amplitude as a function of cone ratio as defined by Equation 8.23, following Gunther and Dobkins (2002)'s definition. We can see the shift to 0.9 due to the ratio of background excitation for L & M cones, see text for details. The scale of the figure has been changed to display the effect of the shift.



(d) Relationship of L and M contrast amplitude as a function of normalized cone ratio, modified by background cone ratio. We can see the shift to 0.9 due to the ratio of background excitation for L & M cones, see text for details. The scale of the figure has been changed to display the effect of the shift.



(e) b and cone ratio. As L and M cone contrast are related by the parameter b in Equation 8.6. This figure is the complementary of Figure 8.3c above.



(f) b and normalized cone ratio. As L and M cone contrast are related by the parameter b in Equation 8.6. This figure is the complementary of Figure 8.3d above.

Figure 8.3: Description of the behaviour of cone ratios, L/M left column and L/L+M right column. These values corresponds to an example in which we have a nominal amplitude of -0.02 for M and +0.02 for L. The values of adjustment (b are therefore restricted within that range).

actual cone numerosity ratio in the retina.

8.2.7 Equiluminance settings

In summary, I have reviewed that the L/M contributions to the luminance channel varies between observers and within the retinal position of each observer. In order to make the least assumptions it is therefore necessary to test equiluminance for every contrast that is going to be presented in conditions similar or as close as possible to the final stimuli.

In practice it is needed to determine b for every value of a to be displayed. As I have shown in previous sections, one can then determine individual cone contrasts (and cone contrasts in other contrast spaces) as well as infer the L/M cone ratio, or cone weighting.

There are two main techniques to obtain these values in order to make stimuli isoluminant. I describe them in the next section.

8.2.8 Flicker and motion perception at isoluminance

As we have seen in the main Physiology review (Chapter 2) the Magnocellular pathway inputs into (somewhat) segregated layers in V1, some of which project into the MT area specialized for motion perception.

This is convenient when one wants to study isoluminance. If it is possible to silence the flicker or motion percept then it is reasonable to assume that there is no luminance modulation feeding into the detection mechanism, hence the stimulus is then isoluminant. Consequently two main techniques have been used to determine isoluminance, heterochromatic flicker photometry (Pokorny, Smith, & Lutze, 1989; Wagner & Boynton, 1972) and minimum motion technique (Anstis & Cavanagh, 1983). Note that other techniques exist such as minimum border perception (Wagner & Boynton, 1972).

For all the techniques describe above, the percept that the experimenter tries to extinguish is never (or rarely) fully nulled. That is why in the descriptions of their respective methodology the goal is always to find the minimum percept.

HTP (heteroflicker photometry) does not require a repetitive pattern and historically has been used as a single field modulating its chromaticity in time. Similarly, the minimum border experiment is usually performed with a bi-partite field. For historical reasons, these techniques were used to determine photopic curves (see for example Guild, 1931, using bi-partite brightness matching, and Methods Chapter 4 on Photometry and colorimetry). As noted by Wagner and Boynton (1972), as the technique used to determine photopic curves were using flicker or correlated measures, the results of nominal isoluminance (see Section 8.2.1) is already quite close to the true isoluminance of one observer albeit not good enough for individual equiluminance.

The minimum motion technique was defined using spatially recurrent stimuli and it is possible to adapt the protocol and stimuli of the HTF technique. The common practice used in the experimental part of this thesis is to make the stimuli used in the isoluminant experiment as close as possible to the one used in the following calibrated experiment. I refer to the modified HTF technique as minimum flicker technique, as the stimuli used differ from what is commonly used in photometric experiments, however the phenomenon

on which this experiment relies is similar in both.

In the two following sections, I describe the data I obtained using minimum motion and minimum flicker techniques. In each section, I describe in more details the methodology and data analysis.

8.3 Minimum motion experiments

8.3.1 Aim

The aim of this experiment was to set the stimuli to be used for the colour shading experiments (Chapter 5). I chose to measure minimum motion for every contrast value (a) presented for each participant and for every pattern type (sinusoid and square-wave). By testing all the conditions presented in the final experiments, the number of assumptions made about equiluminance is limited.

Additionally these experiments will answer two parallel questions. Firstly, this will test the linearity of adjustment (b) with contrast amplitude. If the ratio of contrast from L to M is constant with contrast, then there would be a linear relationship between a and b. The more conservative hypothesis would predict a monotonic relationship or in an extreme case results would show an exponential relationship. Note that the latter hypothesis corresponds to an asymmetric processing of contrast between L and M inputs.

By testing all the contrast values there is no need to make any assumptions about these relationships as every stimulus will be isoluminant by using the values from the minimum-motion experiment directly. Secondly, by testing both square-wave and sinusoids, it is possible to compare the processing of both types of pattern. In square-wave patterns more harmonics are present than in a single sine-wave and it has been demonstrated that the hard-edge pattern can modify processing in itself (Kingdom and Simmons (1998), for isoluminance). Furthermore, the perceived contrast of a sinusoid compared to a square-wave is different (Ginsburg et al., 1980) for similar Michelson contrast. Once again, in this test I make no assumptions as I aim to test both patterns for all contrast values, however the results from such experiment are interesting in their own right in terms of contrast and isoluminance processing of sinusoid versus square-wave patterns.

In summary, the purpose of the minimum motion experiment was to obtain parameter b, which will be different for each participant and is dependent on the amount of contrast modulation amplitude (a, in equation 8.6). Additionally, these results will provide data on linearity of adjustment and contrast and on the differences in processing of square-wave versus sinusoid at equiluminance.

8.3.2 Methodology

The techniques used to find isoluminance usually rely on the poor motion, flicker or spatial perception signal carried by the purely chromatic information. To obtain a measure of b, participants performed a minimum motion experiment (as used by Anstis & Cavanagh,

1983). The minimum motion experiment consists of alternating colour and luminance gratings, with phase offsets, at a visible rate. The phase offsets are designed so that if colour gratings possess a luminance component[6], the alternation of this luminance component with luminance frames will give a perception of coherent motion. For example, if the red stripes on the sinusoids are darker than the green ones, then the motion detection system will match them to the dark bars on the luminance frames and, due to the phase offsets between frames, this will create the perception of coherent motion in a specific direction.

Stimuli

The stimuli were pre-processed to generate smooth transitions when adjusted. A range of stimuli for each contrast amplitude a was pre-generated (0.004, 0.008, 0.012, 0.016, 0.02). This essentially changes the parameter b in the stimuli equation. The stimulus design follows the original description from Anstis and Cavanagh (1983).

The stimulus sequences were composed of four frames in total. Two chromatic frames (pre-generated for each contrast value and for a wide range of adjustment values) and two luminance frames at fixed contrast. The luminance frames need to be suprathreshold to induce motion perception.

The four frames were interleaved so that achromatic and chromatic frames followed each other in the time sequence. In the four frame sequence, the two chromatic frames were in anti-phase (180 degrees out of phase relative to each other) and the two luminance frames were in anti-phase. The luminance and chromatic frames were 90 degrees phase offset in the cycle. Consequently, between each frame there is always a 90 degrees spatial shift of the wave. If the chromatic frames contain no luminance artefact at all (equiluminance), then the luminance channel receives only an anti-phase flickering signal from the stimuli but no coherent motion is seen. However, if there is a luminance artefact in the chromatic frames then a coherent motion will be seen in either direction depending on which stripe (red or green) is lighter or darker. Note that there is an alternative method with no achromatic frames but only chromatic ones spatially offset 90° from each other. Such a method would yield very similar results; however, to my knowledge this has never been formally tested. It is possible that in some publications authors mentioned using the minimum motion technique by Anstis and Cavanagh (1983) when using chromatic frames only.

The stimulus characteristics were made as close as possible to the final chromatic stimuli used in the experiment. They were hard-edge windowed sinusoids or square-waves of 4° diameter, the spatial frequency was 0.75 cpd. The chromatic stimuli in our experiment could be either -45 or +45 degrees orientation in our minimum motion. However, I tested the values at the midpoint between those conditions, that is, vertical. Consequently, in the task observers perceived strictly motion to the right or to the left.

[6]In the rest of this chapter, and in this thesis as a whole, I refer to the luminance components in colour stimuli as luminance artefacts.

Procedure

As the point was to find the equiluminant settings i.e. when red and green stripes are perceived to have equal luminance, this translated to searching for the minimum motion percept. The method of adjustment was used. By turning a response knob (Cambridge Research Systems CB7), the participants added or subtracted equal amounts of contrast to both L and M (effectively updating b in Equation 8.6, 8.7) on the chromatic frames. Participants were asked to turn the response knob in the direction opposite to the perceived motion, until they perceived minimum motion. The minimum motion was perceived just before the motion changed direction at the point of perceptually equal luminance between red and green stripes.

Participants were encouraged to take as much time as necessary to make fine judgements. Participants completed twelve settings at each value of the contrast modulation (a that was tested in the colour shading experiment and of course not necessary for 0 colour contrast). All conditions were randomized.

Sinusoids and square-waves were tested on different sessions; the participants performed the minimum motion associated with the same pattern type before performing the colour-shading experiment. Participants performed the minimum motion experiment before performing the colour shading session.

8.3.3 Results

The results of the 12 trials were averaged for each contrast condition and were used directly to configure the shape-from-shading experiment. The results are presented in Figure 8.4 for 12 observers for both sinusoids and square-waves, using the mean and standard error of the mean (error bars).

The results show that, for each participant the resulting sinusoids and square wave minimum motion adjustment tended to be similar. This can be seen for most participants, the results of each pattern type at similar contrast values tend to overlap (or more precisely the SEM's overlap). This is the case with the exception of participants 9 and 12, who mainly show divergence in the two first contrast values.

From the dataset of isoluminance settings it can be seen that there is a linear relationship between b and a, the intercept being 0. This is expected, as there should be no adjustment in the absence of a stimulus. I consider this point in a following section. The usual practice in the literature is to test for one contrast value and extrapolate for the other values. For example, a linear relationship was implied by Kingdom et al. (e.g. 2005). Here I have tested our participants on all contrasts used in the subsequent colour shading experiment.

Linearity

Regarding the first question on linearity, it seems that the results presented here tend to agree. The hypothesis was that there would be a linear relationship between contrast

Figure 8.4: Results of minimum motion experiment: sinusoid and square-wave conditions. Each panel shows results for a single participant. Sinusoid (black circles) and square-wave (grey squares) patterns, averaged results of 12 tests. Error bars: standard error of the mean (SEM). The results are presented in the form of parameters a and b (See text for details). This data is transformed into L and M cone contrast using 4.6. For an example see Figure 8.2, that shows participant P7 data for sinusoids in term of L (A_L) and M (A_M) cone contrasts. Note that the results for both type of patterns (Sinusoids and square-wave) are very similar within all participants. There are however individual differences in matching between participants, as expected (Gunther & Dobkins, 2002).

modulation (a) and adjustment (b). In fact there was two expectations: firstly, at contrast threshold there should be no adjustment as there is no colour contrast present (hence it is the only contrast condition that was not tested in the minimum motion experiment), so the intercept of the data should go towards 0. Secondly, as contrast increases more adjustment is expected. There is no reason why contribution to the luminance channel from L and M cone is dependent on contrast to such an extent that the relative contrast adjustment would vary. Hence, the second expectation was monotonicity. However, according to our own argument it should be more than monotonic as it should not plateau, hence one would expect at least linear processing and by the same account, it should not increase in strength and become exponential.

The second expected result can be related to the L/M cone ratio section (see Section 8.2.6, particularly on the subject of computation), and this leads to an interesting conclusion. From the data collected it is possible to obtain the cone ratio input to the luminance channel at every single contrast value tested and by extension get a cone ratio estimate as well. Now, the cone ratio is expected to be constant with contrast so a different result than linear would suggest that the cone ratio is dependent on the contrast, suggesting non-linear processing of cone inputs regarding contrast. So a result that would be 'more than linear' (e.g. exponential) would also be surprising and unexpected.

Both expectations regarding linearity and intercept seemed to be broadly matched by all of the observers tested, taking into consideration the noise due to the small number of repeats. If linearity seems to be the rule then it is acceptable to use a high value of contrast and then interpolate the other values of contrast using the same contrast ratio. In this methodology, I took the standpoint to use exactly the values adjusted by the participant. Another method would be to fit a linear curve fixed at intercept 0, using linear regression. This method would have the advantage to give one estimate (with confidence interval) instead of several estimates (with associated uncertainty). Alternatively it is also possible to average all estimates (one for each contrast values), in which case the SEM needs to combine several estimates of variance.

Patterns

In this section, I discuss the difference in sinusoid and square-wave patterns, in the context of the minimum motion experiment. Some discussion on patterns has already been seen in Chapter 5. I will first give a quick summary of the discussion on the patterns then discuss the minimum motion results obtain within this context.

It has been shown by P. Sun and Schofield (2012) that sinusoids and square-wave are processed differently in terms of shape from shading with individual differences and additionally showing the effect of edges on the perception of shape (with square-tooth luminance stimuli). Furthermore, in terms of suprathreshold contrast perception, Ginsburg et al. (1980) showed that there were differences, with square-wave being perceived with more amplitude than sinusoids by a fixed ratio of $4/\pi$. Additionally, Kingdom and

Simmons (1998) showed that square-wave stimuli might be processed differently in the chromatic isoluminance channel than in the luminance channel.

Now coming back to the results in the minimum motion experiment, it can be seen from the results, that the two patterns yielded similar results from all the participants. This is an interesting result on two accounts, in terms of contrast processing and in terms of spatial channels. I detail these two interpretations below.

As I described earlier, the perceived contrast of sinusoids and square-waves of similar Michelson contrast[7] is dissimilar with square-waves tending to have higher perceived contrast (Ginsburg et al., 1980). This effect can be demonstrated when viewing my stimuli see Figure 5.2 p. 96, where stimuli for similar contrast values are shown both for sinusoids and square-wave. This is also the reason why it was surprising not to see more differences between sinusoids and square-wave in terms of overall depth perceived (but as I have pointed out this tends to shift the depth curve leftwards). It is then possible that the stimuli used for square-waves displayed a higher perceived contrast than its sinusoid equivalent, and as such, one should be cautious of putting the two on a single figure. However, from the previous section I have discussed that the results tended to show the linearity of correction in relation to contrast amplitude. If this is the case and if there is a single ratio of contrast perception between sinusoids and square-wave, as suggested by Ginsburg et al. (1980) then the two results for sinusoids and square-wave should diverge. If the ratio of contrast between sinusoid and square is not fixed (as it is the case in some specific parameters in Ginsburg et al., 1980) then more complex diverging behaviour should be expected. Importantly the data from Ginsburg et al. (1980) diverged only at low frequency and low contrast and it was only tested for luminance targets. The equiluminance data shows that the perceived contrast has no effect on the actual luminance needed to make it equiluminant.

The results are also relevant to spatial processing. Some results in the literature[8] (Kingdom & Simmons, 1998) suggest that there is specific processing associated with edges. Here, our results show no fundamental differences in term of low level processing (as both patterns seem to be processed in the same way). This might suggests that differences linked to edges might be due to higher processes. As we have seen for higher processes of shape perception, the results in colour shading did show some differences in processing (Section 5.2.2).

In the next section, I will summarize the results and conclusion obtained from the minimum motion experiment.

8.3.4 Summary

I tested the 12 observes who participated in the colour-shading experiment through a minimum-motion experiment. The aim was to calibrate the stimuli to be equiluminant for each observer. This was tested for both sinusoids and square-wave.

[7] Or other metric based on highest/lowest point versus background

[8] This also includes visual illusions such as the Mach band and the Cornsweet effect

The results provided calibrated values to be displayed in the following testing session. Additionally, the results suggest that the adjustments are linear with contrast, and the ratio of L to M contrast hence constant. The second point is that sinusoids and square-waves tend to show similar adjustment, which as we described earlier is surprising as they have additional harmonics on top of fundamental spatial frequency.

8.4 Minimum flicker experiments

8.4.1 Introduction and Aim

The aim of this experiment was to calibrate the stimuli used in the contrast matching experiment (Chapter 6). As for the previous experiment, the choice was made to test the adjustment necessary to make the stimuli equiluminant (b) for all the contrast amplitudes tested (a). Using the same thinking, the aim was to make as few assumptions as possible by testing all contrast values and by making the equiluminant stimuli as close to the final stimuli as possible.

For the contrast matching experiment, a couple of issues were present that were not with the colour shading experiment. The experiment was only using sinusoids, but due to the number of conditions tested (in the contrast matching experiment), it was necessary to split the test into four sessions of (approximately) one hour. So here as opposed to the colour-shading experiment (see Procedure Section 8.3.2) it was not practical to conduct the equiluminant experiment before every session without increasing drastically the duration (or number) of testing sessions.

Consequently, there is an issue with the problem of consistency of equiluminant perception. The general wisdom in the field is to test as regularly as possible, for example Cass et al. (2009) tested equiluminance before each day of testing for each participant. This is easy with few participants and intensive testing, however in my experiments it was planned to use a large(r) number of participants tested in the course of a week (or less). It was therefore decided to test participants before (for calibration) and after, for test-retest reliability. The aim was to search for shift in the adjustment, if such a shift would be found then we cannot assume that the stimuli were equiluminant during the course of the whole 4 sessions and the data should be discarded as contaminated by artefacts.

Some participants delivered the same results for the minimum motion experiment and the minimum flicker one (with a year or so interval between the two experiments). This gave an interesting opportunity to relate colour-shading and contrast perception, but by doing so, I also gathered data regarding equiluminance.

In the Results section 8.4.3, I first present data from all the participants of the Minimum-flicker experiment only. On a latter section below the Results section, I will compare the results obtained from the same participants tested on two different methodologies (minimum-motion and flicker) at two time periods.

The next section details the methodology used to create the stimuli and test the observers.

8.4.2 Methodology

Minimum flicker experiments are also based on the observation that the colour channel has poor flicker detection. The method used here is based on counter-phased flicker. In this case, in order to get the value of b two colour frames are displayed at a set frame rate.

The two frames being presented 180° out of spatial phase. Essentially the red stripes would turn into green stripes and vice versa. If a luminance difference (i.e. artefact) exists between the red and green stripes (in either direction) then the luminance channel would pick a dark-light modulation out of phase, which elicits a flicker perception.

As for the minimum motion task, the design of this experiment was made so that every contrast amplitude was tested to obtain the b parameter for each observer. In order to find this point the observer is allowed to update b in the amplitude modulation of cone contrast (Equation 8.6).

As noted by Webster and Mollon (1993), as opposed to the classical minimum-motion where the direction of the motion helps the observer to determine when they reached (or passed) the point of minimum motion, it might be more difficult for the flicker experiment to find the mid-point. In their implementation, Webster and Mollon (1993) decided use an alternative forced-choice where the participants selected the least flicker of two intervals. I did not choose to use this method, instead I let the participant freely update the luminance component (or b); this was chosen in order to speed up the experiment and collect more data faster. As I am testing 4 contrast values, it would require full psychometric functions to obtain the minimum point with a 2AFC. One could argue that the method of adjustment might be less accurate than a full psychometric definition. However, for equal number of trials the adjustment method is more efficient as it estimates each time the parameter that we want to extract directly.

In order to perform the minimum flicker task a large range of values must be available so that high magnitude of flicker will be visible at high values of b. By using continuous stimuli presentations observers can compare the magnitude of the flicker perceived continuously.[9]

Observers were instructed to turn the response knob in a direction that would reduce the flicker percept; furthermore they were instructed to go above the minimum flicker and adjust backward from there to have a good idea of what the minimum flicker percept is in that setting.

This is important for two reasons, firstly to get more accurate samples but also because the magnitude of flicker might vary between conditions as the contrast of stimuli between conditions is varied. High contrast stimuli tend to have more perceived flicker and, at the minimum flicker point, there is still a strong perception of flicker. On the contrary, at low contrast values the flicker percept can almost disappear.

Stimuli

For the minimum motion implementation, stimuli were pre-processed to generate a smooth transition when adjusted. I generated a range of stimuli for each contrast amplitude a (0.005, 0.015, 0.025, 0.035), for a range of b values. I used Gaussian windowing on the edges, size of 4°, the spatial frequency was the same (0.75 cpd). The chromatic frames were always oriented at +45 degrees.

[9] Observers are essentially conducting a series of perceptual comparisons between the stimuli before and after turning the response knob. However, continuous stimuli can create adaptation effects.

In order to make the stimuli as close as possible to the one used in the experiment I used similar contrast values, the same orientation size and spatial frequency as the stimuli to be used in the contrast matching experiment.

The stimuli sequences were composed of four frames in total. Two chromatic frames (pre-generated for each contrast values and for a wide range of adjustment values) and two grey frames at fixed contrast (mimicking the frame interleaving of the colour-matching stimulus). Each pair of frames is considered as one frame. The interleaving is equivalent to have half the contrast for two frames. The temporal frequency of the two pairs of frames was 25 Hz. The two chromatic frames were presented out of phase with each other.

Procedure

Observers were instructed to adjust a response knob (Cambridge Research Systems CB7); the participants added or subtracted equal amounts of contrast to both L and M (effectively updating b in Equation 8.6) on the chromatic frames. Participants were asked to turn the response knob in the direction that would minimize the strength of the flicker perception. They were encouraged to go back and forth and to explore the range of values available to do so. When the minimum flicker point was found, observers were asked to press the knob and go to the next trial.

Participants were encouraged to take as much time as necessary to make fine judgements. 22 observers were tested in total, observers M2 to M22 were all naive to the task, and observer M1 was the author. All observers were screened with the Ishihara 32-plates for colour vision deficiency. Participants completed twenty settings at each value of the contrast modulation (a), which were tested in the matching experiment plus one at 0.035 and of course there was no necessary testing for 0 colour contrast. Hence, observers performed 80 randomized trials. All conditions were randomized.

Observers were tested on another occasion on the same protocol, after the contrast matching sessions. Most observers performed the second minimum-flicker experiment within a week, with the exception of observer M1 (the author) who was retested a month afterwards.

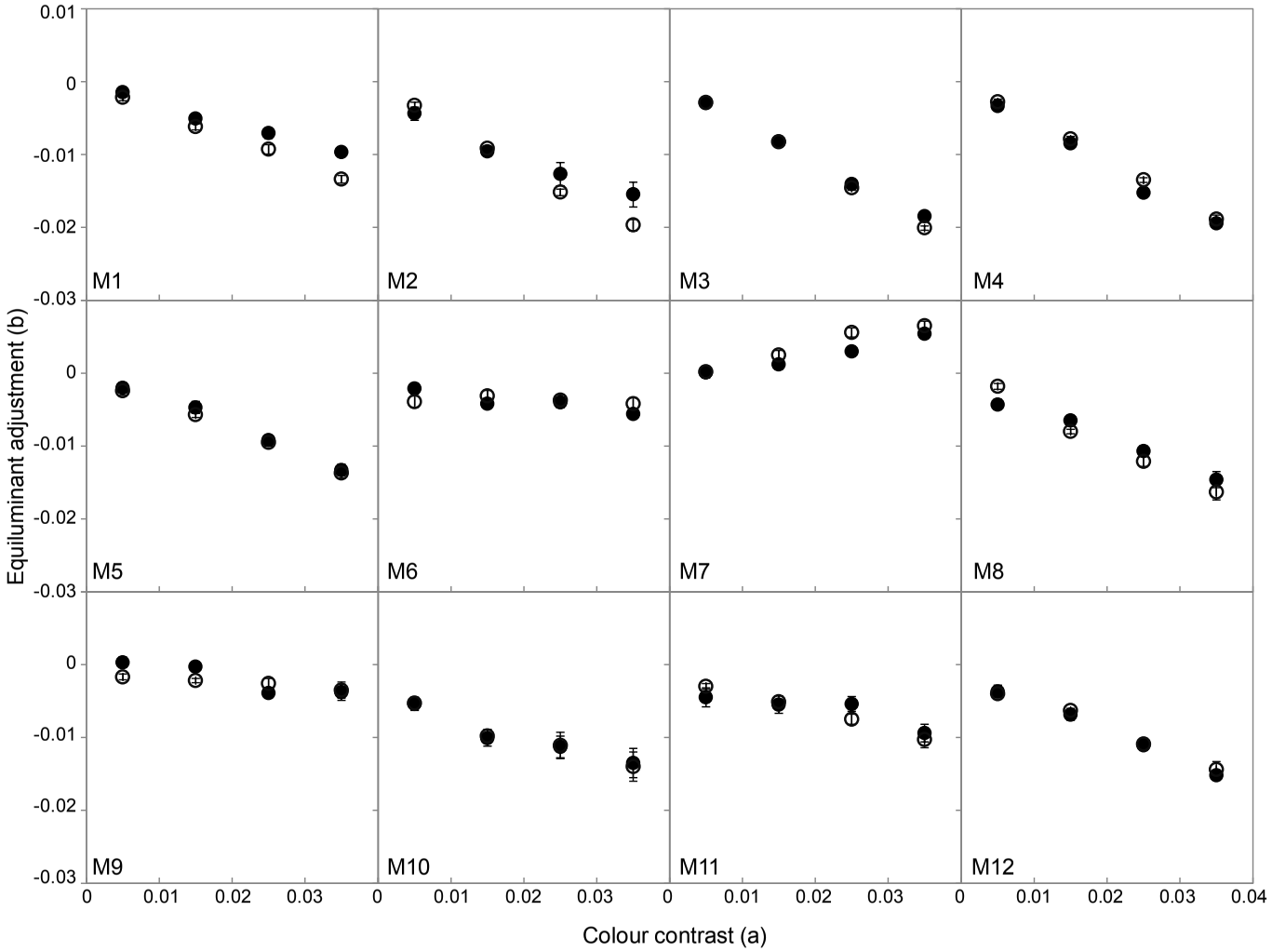


Figure 8.5: Results of minimum flicker experiment for participants M1 to M12. First test (solid circles) and Retest (open circles) averaged results of 20 trials per conditions. Error bars: standard error of the mean (SEM). The results are presented in the form of parameters a and b (See text for details). This data is transformed into L and M cone contrast using (Equation 8.6). Note that the results between test and retest are really close to each other. The linearity of adjustment and individual differences in the matching between participants is notable and as expected from the experiment on minimum motion.

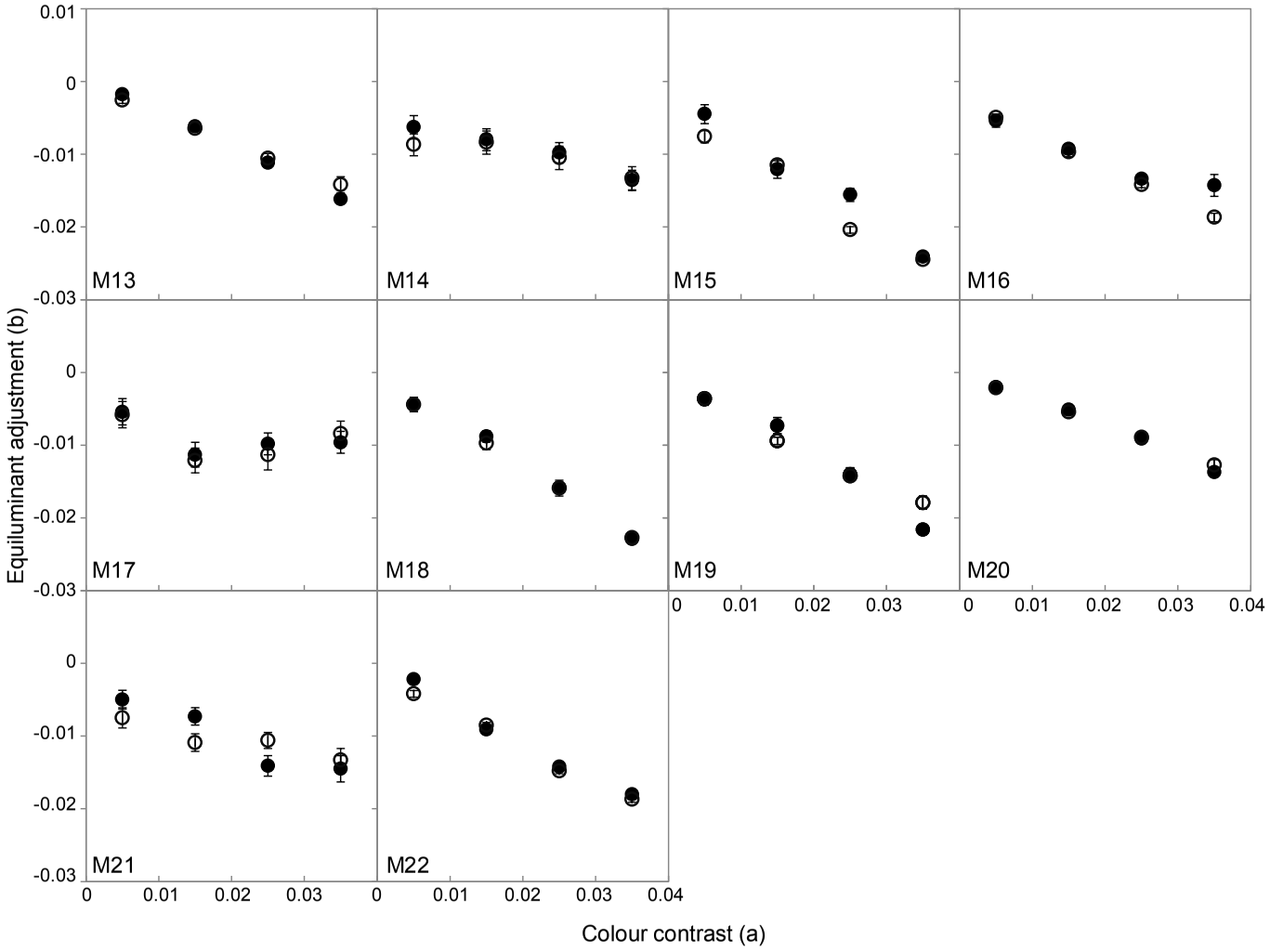


Figure 8.6: Results of minimum flicker experiment for participants M13 to M22. First test (solid circles) and Retest (open circles) averaged results of 20 trials per conditions. Error bars: standard error of the mean (SEM). The results are presented in the form of parameters a and b (See text for details). This data is transformed into L and M cone contrast using (Equation 8.6). Note that the results between test and retest are really close to each other. The linearity of adjustment and individual differences in the matching between participants is notable and as expected from the experiment on minimum motion.

8.4.3 Results

Test-retest

Figure 8.5 and 8.6 show the results obtained for all participants for test (closed symbol) and re-test (open symbol). All results have been separated into two figures for clarity.

For most cases the test and retest error bars overlap, suggesting they were drawn from the same underlying distribution. When test and retest values are non-overlapping it tends to be at the highest contrast value (M1, M2, M15, M18). The highest colour contrast was not tested during the matching experiment (Chapter 6) so I expect the calibration to be sound (although see results for M12 and M14).

This suggests the one-week variability is smaller than the variability of the data obtained during each session (within our number of data points tested).

Confirmation of hypothesis from minimum-motion data

Once again it seems that the results show that the relationship between a and b is linear and with intercept 0. The intercept should be zero because there should be no adjustment at 0 contrast; and that is of course why this value of adjustment is not tested.

This is once again a good argument for using a linear regression and actually representing the adjustment as a line, and the only parameter to fit should be the slope and not the intercept, assumed to be 0. Another way to think about it is the ratio of L to M contrast in the stimuli, which I have shown earlier, can be turned into the cone ratio (see Section 8.2.6).

8.5 Comparison between minimum-motion and minimum-flicker data

In testing the contrast matching experiment, I tried to re-test some observers who participated in the first experiment (Chapter Colour Shading 5). I managed to re-recruit 7 observers (out of 12).

The aim of this section is to compare the results between minimum motion and minimum flicker experiments. Once again, this was not the main aim of these experiments but it is an interesting question in itself. A more systematic exploration of this question would require testing minimum motion and minimum flicker on sessions really close to each other in time. As I showed in the previous section, after a month some fluctuation can appear (see observer M1). The minimum motion and minimum flicker were tested with around a year difference and, of course, using different techniques based on qualitatively different visual attributes, so we would expect some differences.

Figure 8.7 shows the comparison of the data obtained by the two techniques and at the two time points; the data of the first session (out of the test-retest) was displayed for the minimum-flicker data. I found strong similarity between the two techniques, despite

the time delay between the two tests. Nguyen-Tri, Overbury, and Faubert (2003) showed on a large sample (n >200) that there was no systemic relationship between age and equiluminant point. However, it is not known if individual points might change over time for an individual. It would suggest that the methodology used for both experiments is quite robust.

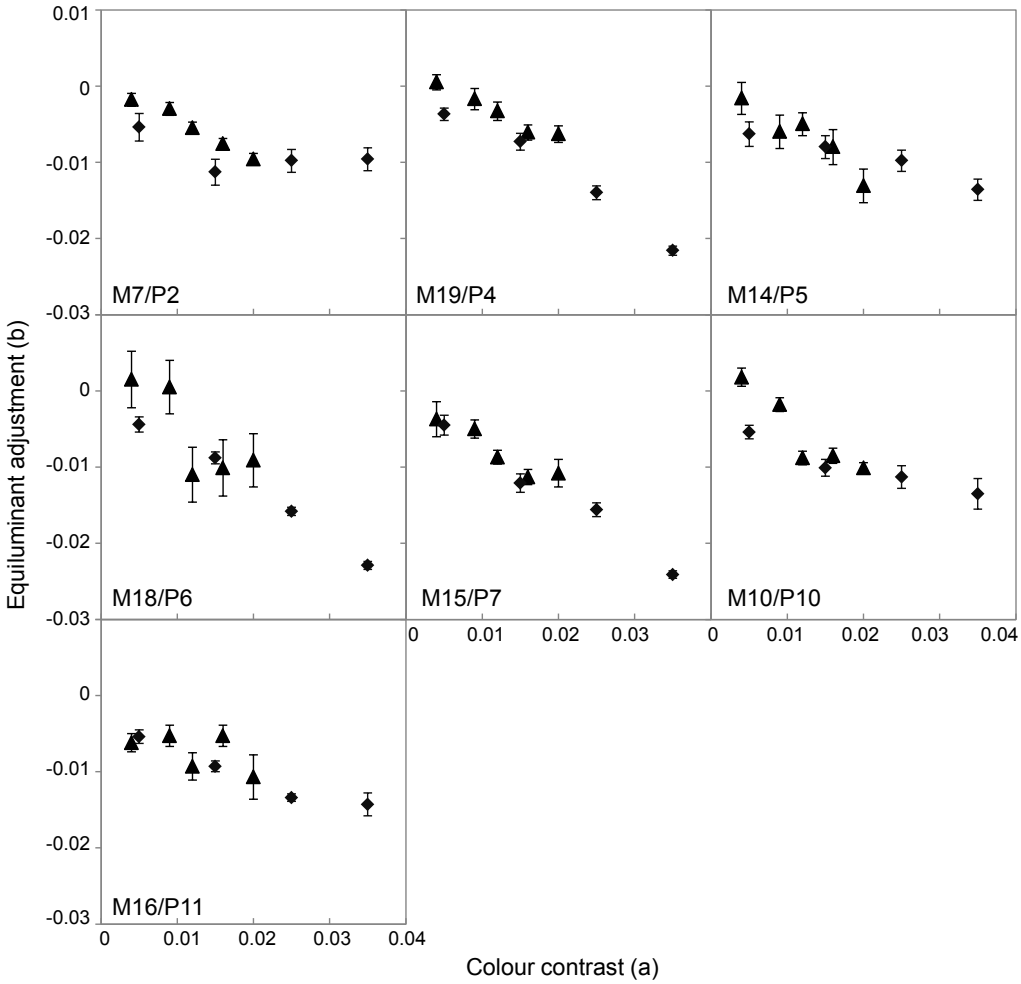


Figure 8.7: Results of participants having been tested in both minimum motion (symbols: triangles), data shown for sinusoids only and minimum flicker (symbols: diamonds). The sessions were performed with over a year in between. Minimum motion data were obtained with 12 data points per contrast and minimum flicker with 20. Error bars: standard error of the mean (SEM). As with previous figures, the results are presented in the form of parameters a and b (See text for details). Part of this data was presented in Figures 8.4, 8.5, 8.6.

8.6 Summary

In this chapter, I have shown the methodology used to calibrate the stimuli for the shape and contrast experiment involving red-green stimuli. The overall aim was to create isoluminant stimuli; however, the results obtained gave relevant insight into isoluminant

processing.

Results tended to be linear, especially in the minimum flicker experiment; hence, the ratio of L to M contrast to generate an isoluminant stimulus is constant with magnitude of contrast. Furthermore, these values are constant between sinusoids and square-waves. I have also shown that observers tend to be very consistent (within a small window of time). I even obtained fairly similar results using flicker and minimum motion over a long period of time (months).

This research opens several questions. Firstly, the variability of equiluminant point with time. Secondly, the difference and similarity of processing of sinusoid and square-wave pattern for motion and flicker. I will conclude this chapter be detailing some insights gained with these experiments regarding three points: linearity of adjustment, specificity of pattern and time fluctuations.

Firstly, from a methodological standpoint I showed that it is possible to test high contrast and assume linear processing to zero. However, a more conservative approach consists of actually testing the exact values beforehand that are going to be tested during the experiment that necessitates calibration. Furthermore, the effect of temporal properties (presentation time, temporal flicker) and spatial properties (spatial frequency, orientation) should be taken into account as it has been shown that they can influence the location of the equiluminance point (see this chapter, Introduction).

On the second methodological point, I have shown that sinusoids and square-wave can be used interchangeably to test equiluminance (in the case of minimum motion). As square-waves are more salient, they can make the task easier for participants. One could potentially use a square-wave of high contrast and work out the adjustment required to make sinusoids of various contrasts (or as suggested before, use the more conservative approach).

On the third point regarding time, it is better to test equiluminance regularly to avoid potential shift. The robustness of the results presented in this Chapter is however quite remarkable.

Chapter 9

General Discussion

In this final chapter, I will summarize the results of this thesis and will discuss them, including shortcomings, methodological points, and proposals for further expansion of the research. So far I have presented data on the colour-shading effect (Chapter 5) on interactions in contrast perception at suprathreshold between luminance and orthogonal components (luminance or colour masks; Chapter 6), and interactions at threshold level between luminance target and orthogonal masks (either luminance or colour; Chapter 7). I have also presented data regarding variability and robustness of isoluminance in my participant population as tested with different methodology (Chapter 8).

9.1 Summary of results

Firstly, I will summarize the main findings related to the topic of this thesis, the interactions between luminance and chromatic signals in visual perception. I will then detail additional findings, not directly related to the main topic but of broader interest in vision science.

9.1.1 Main findings

In this thesis, I have confirmed that modulation of luminance defined shape-from-shading by colour stimuli (i.e. the colour-shading effect) exists and is probably present within a greater variety of stimuli (as demonstrated with the use of square-wave). However, this general point comes with several caveats. The effects observed are dependent on the participant, the type of stimulus (sinusoids or square-wave), and orientation differences between luminance and colour. All participants showed some sort of colour modulation in at least one condition. There is nevertheless no common pattern of behaviour related to chromatic interactions. Furthermore, the data clearly show that colour is not necessary for depth perception (from SFS). However, the effects of luminance contrast were mostly tested in the co-presence of colour.[1]

[1]The baseline condition in condition 2 and 3 when no colour is present is never 0 depth.

The role, if any, of these interactions is still up for debate and will be discussed later in this chapter. Provided that the participants judged the stimuli in the manner expected, there is no evidence for a common heuristic. However, there is the possibility of different heuristics being used (see Section 9.2.4 for a more detailed discussion of heuristics).

The test-retest experiment (Chapter 5) retested two participants on the colour-shading experiments but with the relative phases of luminance and colour changed. This showed no effect of phase (between 0 and 180°). In a detection task paradigm, Gowdy et al. (1999) found no phase dependency at 0 and +180°, but +90° did not elicit facilitation, similarly to Kingdom (2003). However, one study used a detection task (Gowdy et al., 1999) and the other a manipulation of shape-from-shading by colour (colour-shading effect, Kingdom, 2003). This is why a suprathreshold perception test was necessary to evaluate how luminance and colour interact with each other at those levels, without using higher-level percepts such as shape but directly asking for perceived contrast, which is a proxy of intensity variation.

Interactions between luminance and colour have been compared to second order variations in the literature (see for example Kingdom, 2008; Schofield et al., 2006). This is particularly interesting regarding studies performing perceptual training (Dovencioglu et al., 2011). Dovencioglu et al. (2011) showed that the relationship between second-order and first order luminance variations could be learned (even incorrectly). This begs the question whether differences in perceptual learning could explain inter-individual differences and/or differences between trained observers and naïves in colour-luminance interactions? These two questions will be discussed in Section 9.2.5.

Orientation was not studied after the colour-shading experiment (Chapter 5) and studies focused on orthogonal interactions. The following experiment (contrast matching, Chapter 6) showed that perception of luminance contrast, at suprathreshold levels, can be modulated by the presence of chromatic orthogonal components. This is also the case for luminance masks (on luminance targets; see Additional findings, Section 9.1.2). The type of modulation can be either suppressive or enhancing towards the matching contrast.

The chromatic effect observed here is different from the Helmholtz-Kohlrausch effect (Booker, 1981; Burns, Smith, Pokorny, & Elsner, 1982; Wyszecki, 1965), as here the variations of luminance and colour are superimposed, and furthermore, stimuli used are not spot-like but vary around a zero point (sinusoidal). However, there are some similarities (see Section 9.2.2).

The study of luminance vs. chromatic mask effects on the target showed no specific correlation between luminance and chromatic effects. This suggests that no common solution was used for orientation interactions. The complexity of interaction patterns also argues against participants using similar heuristics for both luminance and colour. Instead, a large repertoire of behaviour was found. This also suggests that these results are not due to luminance artifacts in the chromatic stimuli, as they would have produced similar effects to the luminance masks. Effect strength also tended to depend on the target value. No monotonic effect of target is reported for the whole population (although some

participants showed that pattern).

A variety of effects of colour on luminance was found: suppressing, enhancing, or neutral. This matches the variety of effects found for the depth modulation (Chapter 5). The next step to validate the link between contrast effects and colour-shading effects was to check participants who participated in both experiments. However, the results of participants participating in both experiments (a year apart) did not argue strongly in favour of this, as the directions of effects were not all similar between experiments. It is interesting to note, however, that one participant who showed the strongest effect in the colour-shading effect did also show strong luminance modulation afterwards in the contrast matching experiment.

The data of the detection threshold study (Chapter 7) did not show consistent chromatic effects on luminance detection, although some sporadic suppressive effects were found for some participants. No facilitation effects were reported. The lack of strong suppression complements data from Chen et al. (2000a) and Chen et al. (2000b), where suppression only[2] was found for iso-oriented red-green masks on luminance targets. This implies that chromatic-on-luminance suppression is tuned to iso-orientation and becomes weaker at higher values of orientation, with probably some inter-individual differences in the orientation tuning spread. Furthermore, I showed that the lack of facilitation (Chen et al., 2000a), for this type of cross-channel interactions stimuli, is also absent with orientation differences between target and mask.

The facilitations/suppressions observed in suprathreshold contrast might suggest that different mechanisms are at work at those levels (Petrov et al., 2005). A possible mechanism for facilitation is that self-suppression of luminance targets at high values gets inhibited; therefore, the stimuli are boosted. For suppression, it would suggest that the colour component is part of the divisive pool (along with other broadband suppressive mechanisms) of the target. The individual differences between participants might suggest that both mechanisms are present with different weighting between observers. The flexibility and setting of those weights is of particular interest as well as their robustness (see Perceptual learning, Section 9.2.5, Individual differences, Section 9.2.1).

9.1.2 Additional findings

This section reports results not directly linked to luminance and chromatic interactions (the main topic of this thesis) but presents additional results of interest.

Shape-from-shading It appears that depth processing is consistent between sinusoids and square-waves (Chapter 5). The depth perceived is mostly dependent on contrast (Chapter 5, Section 5.2.2), extending results from P. Sun and Schofield (2012). Only one participant could not perceive depth in square-wave stimuli, also coherent with results from P. Sun and Schofield (2012). Perceptual learning could play a role in learning how to

[2]Facilitation can be found when the chromatic stimuli are not properly isoluminant, see Chen et al. (2000a).

interpret the square-wave stimulus (see Section 9.2.5). There might be a stronger prior (or heuristic) for smooth edges than sharp-edges for the few participants unable to perform this task.

Suprathreshold luminance-luminance interactions Suppressing and enhancing effects were found (contrast matching experiment, Chapter 6); this matches and broadens the data obtained by Meese and Hess (2004). It is worth noting that Meese and Hess (2004) only found facilitation in the binocular case (my contrast matching experiment was only using binocular viewing). It would be interesting to see if the effects found in this thesis disappear when viewed monocularly or dichoptically, as expected from the two participants' data in Meese and Hess (2004). Testing various viewing conditions using colour components would also be of interest.

Further, the presence of facilitatory effects in Meese and Hess (2004) was dependent upon the participant and the SF used. The SF dependency is of interest. If these facilitations are not found broadly for all SFs, they might not play an important role for spatial processing or their role is limited to a certain tuned range. Alternatively, it is possible that these facilitations can be tuned specifically to different SFs between participants. Furthermore, inter-SF channel interactions would be interesting to study; facilitation vs. suppression mechanisms should be tuned to the SF distance between mask and test, as it is possibly the case for orientation differences (see main results, Section 9.1.1).

Detection luminance-luminance interactions Once again, the additional findings are concerned with luminance-luminance interactions, but this time for the detection study (Chapter 7), as they were added as both an artifact check and to compare cross-channel interactions with interactions within luminance channels but between orientation channels.

My results, in Chapter 7, add to some evidence in the literature, that at low SF, within-channel interactions have a dipper-function shape. This entails a facilitatory component and a suppressive handle on the TvC curve. My data show four participants with this type of behaviour, with more or less accentuated dipper shapes. This is in accordance with data by Meese and Holmes (2007). Interestingly, they showed that the facilitatory component is reduced when using faster temporal frequency presentations. This shows that facilitatory and suppressing mechanisms are of different nature. The reason for their separate tuning (in terms of orientation, SF, temporal tuning) is not currently known in terms of the computational problem being solved.

9.2 Discussion points

In this section, I will further detail discussion points resulting from the experimental findings presented above. Unlike the previous section, the points below are meant to reflect on the results of the whole thesis and not of specific experiments.[3]

[3]Unless stated differently in text.

9.2.1 Individual differences

One common feature of the results in this thesis is the widespread presence of individual differences. This might usually be a dreaded result but as I will detail below it might not be so. Firstly, I will try to evaluate the source of these variations. Variations between observers might come from several sources; in the worst case it can come from uncontrolled variations in the setup over time or variations in the level of understanding of the task by different participants.

The first critical source of variations mentioned is due to the experimental setup. Variations of the setup over time influencing systematically the results are however unlikely. The setup was calibrated on a regular basis. Furthermore, due to the length of these experiments, participants' sessions were interleaved over time. The second critical source of variation is failure on the participants' side, having variation in the understanding of the task. Regarding improper understanding of the task, there would be indicative signs in the data that are not present. Most of my experiments include baseline conditions, in which case it is easy to see if inappropriate responses were made. This is especially the case when the stimuli are luminance-luminance defined and the participants might respond to the wrong orientation component. However, the experiments consistently asked to follow one orientation and in the most ambiguous case (luminance-luminance in suprathreshold matching), a matching stimulus was always present in the appropriate orientation. Thus, the two majorly problematic reasons for variations can be excluded.

As compared to other studies in visual perception, my sample sizes tended to be larger. Thus, if variation is present in the population, by testing more individuals, I have shown what would have been otherwise overlooked. It is possible that by testing a wider number of participants, I am uncovering what was hidden before by the use of small samples usually composed of the main authors with the occasional naïve observer.

Another reason for the variations might be the use of higher-level tasks, in the depth (from shading) and contrast perception experiments. These two experiments involve processing that requires already some understanding of the 'object' (i.e. stimuli). Scene understanding will be discussed in Section 9.2.3. Variations in the interpretation of the shape might create some of the observed variation.

Strategy One could argue that (some) participants are aware of the modulatory effects of the masks and, despite the instructions to match stimuli and matching probe perceptually, are compensating for this effect. In a related experimental paradigm (see Section 9.2.2), Guth, Donley, and Marrocco (1969) commented on the suggestion that participant might compensate, "Within this framework, to suggest or even imply that subjects must subtract something from their brightness judgement to allow for saturation is defeating the very purpose of an experiment which is designed to obtain psychophysical judgements of stimulus magnitude." (Guth et al., 1969, p. 566).

In this thesis, the difference between strategy and heuristics is that I consider strat-

egy to be a conscious process and heuristics to be an unconscious processing 'strategy'; consequently, heuristics are treated in another section (Section 9.2.4).

Evidence for similarities I will now discuss the data showing participants coherent behaviour providing evidence that participants were performing the task properly. There are three diagnostic indicators of reliability: test-retest, consistency of the data (e.g. baseline matches expectation), and coherence with the literature.

The *test-retest* is found in some cases by design and some other cases it emerges from other requirements (e.g. checking for artifacts). In the first experiment (Chapter 5), some participants were retested with a change of phase to check for artefacts. Similar results were obtained. Furthermore, for the whole dataset sinusoid and square-wave data tended to be similar for the equiluminance adjustment (Chapter 8). Furthermore, the luminance effect (shape- from-shading) showed strongly the same shape for sinusoids and square-waves (Chapter 5). Although not properly a test-retest, this shows the robustness of the effect. In the second experiment pre- and post-equiluminance tests were very similar. Interestingly, participants tested in the colour shading experiment showed similar results for the adjustment using a different adjustment technique with a significant time between tests and on a different experimental setup (compare Chapter 5 & 6).

Consistency in the data corresponds to the data having meaningful values in regard to the physical stimulus and the percept, for example no negative depth perceived or no contrast perceived when there is no target, monotonicity of some effects and so on. In the first two experiments no-target stimuli trials were added as catch trials (e.g. 0 contrast). This showed good understanding by the participants. There were also baseline conditions used (e.g. Chapter 6), where no interfering masks were present, to check the reliability of the participant and get a baseline value in case of biases in the responses. Experiment 2 (Chapter 6) was rich in target conditions and therefore had a large number of baseline conditions, which showed that the participants performed the task correctly (see Section 6.2.3). Furthermore, consistency can consist of the lack of adaptive effects or drifts in the data. This was checked for the contrast matching experiment where such effects can occur, and no evidence of adaptation was found (see Section 6.2.3).

Coherence with the literature means that the data are consistent with what is currently known about the visual system. The monotonicity of shape-from-shading with luminance contrast is a good example. Furthermore, the monotonicity of the adjustment with colour contrast is a good indicator that the stimuli were adjusted properly. In terms of luminance processing, the shape-from-shading results (Chapter 5) and the dipper-shaper results (Chapter 7) are good matches with the literature (e.g. Meese & Holmes, 2007).

Conclusion I have discussed some reasons why the individual differences were found in the results. Additionally, I have detailed a set of evidence that shows the robustness of the data (test-retest, consistency within data, and coherence with the literature). The individual differences are likely due to differences in visual processing rather than experimental

shortcomings.

9.2.2 Helmholtz-Kohlrausch effect

In Chapter 6, on contrast perception modulation, I compared the effects found for the contrast perception experiment to the Helmholtz-Kohlrausch effect (HK effect). In this section, I will discuss further the similarities and implications.

Definition

The HK effect corresponds to changes in brightness perception[4] linked to colour saturation/contrast (Booker, 1981; Burns et al., 1982; Wyszecki, 1965). The phenomenon also refers to monochromatic light having less perceived brightness than white light of similar luminance, which has been called failure of additivity (Guth et al., 1969). This is a failure of additivity between colours to drive brightness perception. Sometimes the additivity can also be greater than the parts (Guth et al., 1969). This effect is greater with saturation (contrast), see for example Booker (1981).

In a series of papers, Sanders and Wyszecki (1957) and Wyszecki and Sanders (1957), using perceived brightness of objects,[5] found that luminance was an inaccurate proxy of perceived brightness. This is also the case with lights (Booker, 1981; Burns et al., 1982). It is a failure of the CIE system to match perception in terms of light brightness (Sanders & Wyszecki, 1957).

Modulation of the HK effect

The background value does affect the brightness match (Wyszecki & Sanders, 1957). The background effect works for colour (e.g., red, yellow), but differences in the intensity of neutral backgrounds also shape the effect (i.e. white vs. black background). Wyszecki and Sanders (1957) concluded that saturation was modified by the lightness of the background values and is linked to differential adaptation of the eye.

Sanders and Wyszecki (1958) performed the same experiment as in Sanders and Wyszecki (1957) and Wyszecki and Sanders (1957). They noted that brightness perception "may change appreciably over a longer period". This is interesting in terms of modularity of lightness/colour perception and in terms of perceptual plasticity (see Section 9.2.5, for a discussion of plasticity).

Colour and size of the stimuli are also changing the effect (e.g. Booker, 1981; Sanders & Wyszecki, 1957). Interestingly, Sanders and Wyszecki (1957) noted that the blue cone contribution to the luminance percept decreased with the size of the stimulus (see Physiology (Chapter 2), cone sampling of blue cones in the periphery). Booker (1981) hypothesized that results regarding red stimuli of variable size were due to the difference in L versus M cone sampling from fovea to periphery. Burns et al. (1982) confirmed that

[4]Perceived luminance or psychological correlates of luminance.
[5]Brightness of objects was termed lightness.

the highest effects were found with (monochromatic) blue and red colours. It seems that colour is integrated over a larger area and, therefore, contributes more to the HK effect with larger stimuli (Booker, 1981). Spatial integration is usually overlooked in spatial vision. One of the "dogmas" in this literature, according to Meese and Baker (2011), is that spatial pooling only exists at threshold levels. It would be interesting to change the size of the stimuli used in my experiment to see if the spatial integration has an effect on cross-channel interactions (see Spatial content section, Section 9.2.3 and Section 9.2.6).

It should be noted that using a yellow versus a black (or white) background produces strong differences in the equi-brightness lines on the CIE diagram. It would be interesting to change the background (colour & intensity) in some of the experiments I used. Note that Kingdom (2003), used a yellow/brown background in his original paper. It is possible that this background produces more chromatic effects than the grey ones used in Kingdom et al. (2005), Kingdom et al. (2006), and in my experiments.[6]

Individual differences

It is possible to represent equal-brightness on the CIE chart by plotting equal brightness lines. There are individual differences in those equi-brightness line on the CIE diagram (Burns et al., 1982; Sanders & Wyszecki, 1957; Wyszecki & Sanders, 1957); however, the shapes are similar (Burns et al., 1982).[7] Booker (1981) did find that higher magnitude leads to higher variations in the effects (between participants). This is an interesting parallel to the divergence of effects with higher contrasts in my matching task (See Section 9.2.1, on individual differences).

A complex combination of channels

Guth et al. (1969) outlined that, as Hering (1964) suggested, brightness intensity might be due not only to achromatic channels but also to "some function of responses from both an achromatic and a chromatic system" (Guth et al., 1969, p. 566).

Other percepts, by opposition, are assumed to work only using luminance variations; this is the case for flicker perception. Therefore, flicker perception is only based on the input from one channel and that is why it is used to set-up equiluminance (see Methods Chapter 4 & Isoluminance Chapter 8).

Burns et al. (1982) reviewed possible models of HK effects where brightness was combined from several channels; however, they stated that "Either the initial assumptions of channel linearity are wrong or the combination rule is more complex" (Burns et al., 1982, p. 1230).

There might not be a brightness system in the brain but instead these percepts might

[6] These all used the same grey background values.

[7] Investigation by the Committee on Uniform Color Scales of the Optical Society of America and reported by Wyszecki (1967) looked into an impressive sample size of 76 participants on brightness perception to produce an average standard.

correspond to the intensity estimate of the chromatic and luminance system. This estimate, as pointed out above, might be complex rather than simply additive.

Interestingly, during a pilot experiment for the contrast experiment (Chapter 6), I used a condition with aligned colour and luminance components[8] and asked participants to extract the luminance contrast. This task turned out to be very difficult, to the point that the aligned condition was not used in the final design of the contrast matching experiment. Distinguishing between luminance and colour when aligned seems to be almost impossible without 'expert' knowledge of the stimulus. Therefore, co-linearity of variations makes the sources of intensity variation difficult to interpret/extract. The role of spatial content will be discussed in a following section (Section 9.2.3; see also Section 9.2.6).

Conclusion

There are many similarities between the HK effect and the results found in this thesis. However, one major point sets them apart, namely the role of spatial content. The HK effect is solely studied as a uniform patch comparison paradigm, whereas the experiments presented in this thesis have modulations of chromatic and luminance content within the stimuli (uniform vs. patch).

9.2.3 Spatial processing

The set of experiments presented showed that orthogonal components (luminance or chromatic) interfere with the processing of differently oriented luminance modulations. This has some implications for edge perception, smooth surface curvature, and shading perception. In the colour-shading experiment (Chapter 5), the depth measured was curvature depth modulation of the object. In terms of shading, the contrast matching experiment can be thought of as an object shading matching. The particular point of object perception/understanding will be discussed in the following subsection.

Scene understanding shape vs. contrast

The stimuli used in my experiments do not match real world objects. This prevents the use of extra knowledge (cues) about real world objects that could interfere in the perceptual interpretation. For example, it is clear that knowledge of the colour of objects can influence the perception of colour, as commonly tested with colour constancy (e.g. Granzier & Gegenfurtner, 2012). Another example is the presence of object edges; these can be strong indicators of shape. This in turn can be changed by visual experience (see Section 9.2.5). Furthermore, higher detection accuracy is also obtained in isoluminant objects vs. non-objects (Martinovic, Mordal, & Wuerger, 2011), showing once again that object perception is special.

In the Contrast Matching Chapter (Section 6.3.3), I discussed the possibility of interactions between shape and contrast. Indeed, it is likely that perception modules, for

[8]This was done to fully match the stimuli used in the colour-shading experiment.

shape and shading of a surface, interact with each other (see Psychophysics introduction on lightness and shape interactions, Section 3.6.5, p. 63). However, it is not certain that such dedicated modules exist (see Discussion point on emerging properties in Section 9.2.5).

One of my hypotheses for the colour-shading effect was that the luminance shading, modulated by colour, modulates in turn the shape perception. However, the opposite effect, shape influencing brightness/luminance, can also occur (Section 3.6.5). The final percept is the result of these two interacting process, as postulated by Adelson and Pentland's (1996) model. In the data presented in Chapter 6, evidence for both accounts was found. Adelson and Pentland's (1996) computational model suggests that a visual system needs to extract shading, shadow, and structure formally to understand the scene properly. Olmos and Kingdom (2004) did present a model incorporating colour input to work out reflectance and shading (which could be enhanced to integrate shape as in Adelson & Pentland, 1996).

It is probable that the modulatory effects on shape curvature (modulation depth, colour shading experiment, Chapter 5) and shading (or contrast perception, Chapter 6) found in this thesis would vary using stimuli designed to resemble real life objects. There is evidence that objects are processed faster when luminance defined than chromatically defined, showing a processing advantage for the former (Martinovic et al., 2011). Furthermore, the scene understanding and not only the object itself can play a role. Placing our stimuli in an environment displaying light, shadows, and stereopsis depth would certainly trigger higher processing, influencing the perception of the object as a contextual effect. There are several cues that can be interpreted (depth, size of object, brightness of the object) to work out the true reflectance of an object. These different strategies, or more likely heuristics, will vary strongly with experience (see Section Heuristics 9.2.4 and Perceptual learning 9.2.5) and available information (contextual modulation).

Remarks on the computational approach

In this section, I will discuss theoretical issues relating to the computational approach to study visual phenomena, such as the ones described in this thesis. Some of these issues come from the limited scope of some computational models.

Adelson and Pentland (1996) tried to tackle some of the limitations of the assumptions of shape-from-shading (uniform reflectance and variations in shading only from the shape). The complexity of the problem comes from the plurality of sources that can create luminance variations. They used non-independent analysis modules which matched psychophysical results (see above shape and brightness interactions, Section 9.2.3). Their model ignores colour variations, but in terms of performance, it goes in the right direction. A good example of it is the reversal of brightness perception with reversal of shape perception obtained with their model (Adelson & Pentland, 1996).

I see one conceptual issue that is present with vision algorithms. Even though there is

a tendency to integrate more and more biologically possible computations, which is a good feature for models trying to understand the human visual system (or that of other related species), there seem to be some key biological constraints missing. Some of these missing features were pointed out by Ramachandran (1991). Ramachandran (1991), discussed the transitions between layers in Marr's (1982) theory. Ramachandran (1991) emphasized constraints when implementing solutions, suggesting that a problem should be studied at different levels concurrently.

Furthermore, the goal of these models, which Marr (1982) would have called the computational question, or the problem to be solved might be the fundamental issue. Is the role (or one of the roles) of the visual system to extract reflectance and shading and shadows? This is admittedly a very difficult question. It is acceptable to design an algorithm for computer vision whose sole goals are the ones detailed above and to be clear the authors do not make assertions that this is how the visual system is actually working.

The interesting question here lies with embedded/autonomous systems. The overall goal of the VS according to Marr (1982) was to extract meaningful regularities in the visual world statistics. However, autonomous agents have added requirements. These autonomous systems require visual processing to function and navigate within their environment. The processing has to be robust and efficient and avoid waste in energy consumption. It seems to a certain extent that the human visual system (VS) is implementing such computation requirements, as the encoding of information tries to avoid redundancy (e.g. Hansen & Gegenfurtner, 2009) and coding efficiency is important. The trade-off is important here and, in some cases, 'just good enough' solutions can be sufficient to solve a problem (Kingdom et al., 2005; Ramachandran, 1991). Furthermore, Ramachandran (1991), made an extra point that we should take the evolutionary process into account, since the visual system is constrained by its predecessors. These are two biological points that tend to be overlooked.

This leads me to several conclusions regarding that subject. There is probably no shadow and reflectance centre in the visual system. However, humans are capable of extracting such information. If the detection of such characteristics were vital to survival (and had been for generations), one might expect to find dedicated processing to have its dedicated area (or evolutionary encoded heuristics as suggest by Ramachandran, 1991, see Section 9.2.4). The different characteristics of a scene inform each other; shape, brightness, and other attributes interact with each other. The extent to which these interactions are 'pre-programmed' is not known but adaptability makes perception more versatile and useful to the agent. Additionally, it would be useful to have heuristics or perceptual shortcuts to process/extract high-level information with simple computational costs (see Heuristics section 9.2.4). Extraction of information should not be performed in two separated modules but should inform the other percept, as suggested by Adelson and Pentland (1996). Studying one perceptual feature alone provides valuable information towards visual processing, but we should not forget the greater picture as the perception of such percepts (e.g., shape, contrast) is not done independently. Furthermore, some information

can be emerging from computation without it being the primary goal. For example Meese and Baker (2011) described the spatial summation as an emergent property of dichoptic stimulation with no special role.

9.2.4 Heuristics

I previously mentioned some discussion points on heuristics in the computational section. Kingdom (2003) suggested that the reason luminance and colour interact with each other in the colour-shading effect was the use of heuristics to determine reflectance and shape. This is useful in order to solve the problem of variation due to material light properties versus material shape. In this section, I will discuss if my data could be interpreted as due to heuristics. First, I will present the heuristics hypothesis through the utilitarian theory of perception as suggested by Ramachandran (1991).

Utilitarian theory of perception

Ramachandran (1991) suggested several points the AI approach (Marr, 1982) is missing, one of which has been described before (see Section 9.2.3). These points revolve around neglected biological constraints. Ramachandran (1991) suggested that the implementation of a solution might be different in actual biological systems. Furthermore, he also suggested that we should not ignore the evolutionary history of a perceptual system. In a sense, the biological substrate is both a constraint on the possible computations and already a solution of past 'solved' problems. Ramachandran (1991) also pointed out that perception does not arise from solving a set of equations or internal representations but through a 'bag of tricks' (i.e. heuristics) which are computational shortcuts to complex problems and these solutions appeared through evolution.

Some of these points were applied to the section on the computational approach (Section 9.2.3). One of the problems that I mentioned was about the computational problem being solved. For Ramachandran (1991) the role of the visual system is to determine the general spatial layout, although the authors also acknowledge that if perception is only done through tricks, it is difficult to see how it can create a rich and clear visual percept. In comparison, for Marr (1982) the role of the VS is to extract regularities in visual statistics for a specific task (Kingdom, 2008; Marr, 1982).

The strength of the utilitarian theory is that the percept is created through the parallel use of a lot of 'tricks' and not only one solution, therefore creating robustness. The reason for this bag of tricks instead of a whole strategy is, according to Ramachandran (1991), the lack of foresight in the evolutionary process. In addition, it permits 'good enough for the job' strategies to be used, which do not need to be optimized perfectly as other solutions can be combined with it. This is interesting in the context of this thesis, as it has been suggested (Kingdom, 2003) that some of the reported effects (e.g., Colour-shading, 5) are due to the use of heuristics by the VS to solve a specific computational question.

Colour-luminance interactions as heuristics

Kingdom (2008) outlined a long list of heuristics that the visual system can use to determine light versus reflectance. Kingdom (2008) also reviewed the Bayesian hypothesis where the prior knowledge is encoded explicitly and used against visual input to make perceptual decisions. The Bayesian approach to perception tends to be popular nowadays, however, the implications in terms of implementations tend to be understated and there is not sufficient physiological evidence to back it up. In terms of implementation the VS needs to have a probability distribution over a large scene and a large number of dimensions. In Kingdom (2008), communications from Mamassian are reported where he suggested that a Bayesian framework could potentially explain the colour-shading effect.

The data I showed in Chapter 5 & 6, shows a lot of diversity in the way colour and luminance are combined. In itself, this does not suggest the absence of heuristics; it could suggest a diversity of them. This conclusion works, assuming all participants performed the tasks similarly with no extra heuristics at play but other perceptual modules could interact (see Individual difference Section 9.2.1). This points towards one problem, in my opinion, of Ramachandran's (1991) theory of utilitarian perception. His theory acknowledges biological facts usually ignored by other perceptual theories but neglects flexibility, training, or learning in the system, which are qualities demonstrated by our VS. Ramachandran (1991) takes the example of low-level heuristics that applies to animals.[9] However, we cannot assume that all visual capabilities are derived through evolution, as there has to be a rule for perceptual learning too. I suggest that some heuristics could be added or learned later on.

A modified hypothesis is that these interactions emerge from purely extracting statistics of the visual environment (Marr, 1982) but not necessary through driven goals. In the case of my data, these interactions would be present because of the statistics experienced by observers but not by the purely goal driven design of finding reflectance vs. illuminations (for another example of emerging properties, e.g. spatial pooling, see Meese & Baker, 2011). To be more precise, some important fundamental computations (or more precisely its physiological hardware) might still be encoded in the genes but not to the extent described by Ramachandran (1991), allowing for new rules[10] to be acquired through perceptual learning.

9.2.5 Perceptual learning and visual experience

If the role of the visual system is to extract regularities in the visual environment (Marr's (1982) first stage), then perceptual learning is the process of updating or finding new regularities/rules.

[9] For example, Ramachandran (1991) mentioned the sea-gull infant recognized a parent only through the use of a shortcut, recognizing the red circle on their beak. Therefore, a complex set of behaviours is started with minimal visual processing. A perfect example of perceptual heuristics.

[10] These rules can be heuristics or statistics of the visual environment.

Layers of plasticity

The human brain is often praised for its plasticity. Perceptual learning is one of those examples where our vision can adapt to changes in the environment. Perceptual learning is one of the forms of adaptation to visual statistics. Some other forms include temporary adaptation (see Psychophysics literature review Section 3.2), such as chromatic adaptation, whose functional role has been linked to discounting illumination. Therefore, there is embedded knowledge in the VS that the colour environment can change and our VS can change the perception accordingly to discard illumination and work out object properties.

Some of the modularity of the VS is dependent on the physiology directly, such as receptor adaptation or bleaching and neuron adaptation (Section 3.2). Furthermore, the effects of neuromodulators (such as Acetylcholine, linked to attentional effects) can modify the behaviour (e.g., receptor field and other tuning properties) and dynamics of neuron responses. Changes in the behaviour of these mechanisms might require long-term evolutionary processes. Some rules of processing are due to the setting up of connections during the so-called critical period. The setup of connections usually requires a mix of visual inputs and neurogenesis effects. The visual inputs 'configure' the system using statistics of the visual environment during the critical period. For example, this critical period seems to be important to wire binocular responses (requiring proper alignment of the eyes).

Perceptual learning and luminance colour-interactions

It is required that a higher-level adaptation mechanism exists, allowing the learning of complex stimuli configurations or the updating of learned associations. The goal of this layer of plasticity is to adapt to more prolonged changes in statistics (i.e. longer than neuronal dynamics adaptation). This might require longer changes in synaptic weights or new connections altogether. I would suggest that most heuristics discussed above actually fall into that category. For example, there is no direct need to encode genetically the 'light from above' prior (in shape-from-shading) if it can be learned through visual experience. The learning of this prior would just be the natural consequence of living in such an environment. One could argue that it is not of survival importance to encode such processing. Coming back to the colour-shading or suprathreshold contrast, I would hypothesize that those types of interaction might be due to this higher level of perceptual learning. Furthermore, it is possible that those types of interaction do not require a functional role; the functional role would be an emerging property of encoding visual statistics. One particular process plays a functional role if it adds a benefit to the system. In this case, for some of my participants these interactions can be helpful to work out reflectance versus shape. Other processing might be detrimental for that type of problem (i.e. reflectance versus shape).

Interestingly, Dovencioglu et al. (2013) used perceptual training to 'teach' participants to use second order cues for layer decompositions. It is interesting because second order cues have been compared to chromatic information in their interactions with shape-from-

shading. It is especially worth noting that one of the participants in their study did learn the opposite rule than intended.

Trained and naïve observers

I briefly discussed the point of the use of trained observers in Chapter 5. There are usually three categories of participants reported in psychophysics experiments with different levels of experience with experiments and of knowledge about the stimuli, the task, and the objective. These are naïve observers, trained observers, and authors. However, the first two categories can overlap.

An interesting category of participants is expert observers or expert viewers. This term is usually used to signify that these observers have had intensive training in performing specific visual tasks (common examples are painters, photographers, pilots, or athletes requiring fast visual integration). This once again is a testimony that the VS can improve performance drastically with practice.

The difference between trained observer and author knowledge is in the amount of knowledge about the stimuli. However, any difference in the results would be difference in response strategy to compensate for a possible known effect, and should not happen.

Regarding the data in this thesis, it would be interesting to use expert viewers who have a strong appreciation of reflectance and shape changes. However, it is difficult to know where the difference in behaviour lies, heuristics or strategies. It might be possible to differentiate the two by using passive viewing (prolonged stimulus exposure) vs. guided learning with feedback.

9.2.6 Stimuli

In this section, I will discuss the stimuli used in this thesis in relation to low level spatial processing ((spatial pooling) and (sinusoids vs. square-waves differences)). I will then suggest other stimuli for further studies of complementary interest.

Spatial pooling There are two distinct suppression mechanisms (DeAngelis et al., 1994, 1992), overlay and surround (see Section 2.4.3, 2.4.4 & 2.4.5). With large stimuli, we cannot predict how the two mechanisms will interact (Petrov et al., 2005). Spatial pooling at suprathreshold levels tends to be overlooked (Meese & Baker, 2011). Furthermore, we need to know how the pooling mechanism works to produce conscious perception (at suprathreshold levels). Petrov et al. (2005) noted that the transition from detection to suprathreshold levels is not clear and this is where contrast perception experiments are important, as the study of perceptual channels should not be restricted to detection only.

In this thesis, I did try to compare detection and suprathreshold tasks (Chapter 6 & 7), however, it seems that both target and mask contrast levels influence the amount of contrast in the target (this was clearly the case in the contrast matching chapter, Chapter 6). The study of pooling mechanisms should be performed both at threshold

and at suprathreshold levels. This modulatory effect have been shown in the physiology chapter (See Section 2.4.3). These different characteristics should be studied together. As Benardete and Kaplan (1999) showed, gain and tuning of M cells could be changed by spatial content and contrast. In conclusion, I would suggest studying the spatial pooling phenomenon and unifying the data between detection and suprathreshold levels as well as the physiology. Cross-channel interactions between chromatic and luminance (in the context of spatial pooling) would be of value in the literature. Furthermore, one area of interest is to figure out how population encoding (of neurons) translates to perception of hue, saturation and brightness (see Shapley & Hawken, 2011).

Sinusoids vs. square-waves The point of the addition of square-wave stimuli in the first experiment was to investigate the effect of edges while simultaneously conserving spatial frequency[11] and contrast.[12]

The underlying point was that in our visual environment sharp edges and smooth edges are associated with variations in reflectance or illumination, respectively (Horn, 1974). This was of particular significance for the colour-shading experiment because the heuristics associated with it (Kingdom, 2003) were hypothesized to do this very computation. However, no specific differences were found between sinusoid and square-waves. Interestingly, using prior knowledge one should not interpret square-wave luminance as a corrugated material but strongly interpret reflectance changes (Horn, 1974). This is possibly what one observer (Chapter 5) was interpreting as this observer was unable to perceive any depth in the stimulus.

It would also be an interesting addition to the contrast matching experiment to test square-waves. The purpose of this would be to see if there is a differential treatment of the pattern types (i.e. sinusoids vs. square-waves). If present, this could be linked to differential processing of different sources of variation (reflectance vs. illumination).

Furthermore, when aligned, stimuli might be combined more strongly (see Section 9.2.2, difficulty to extract luminance information in pilot study). In this case, a design with aligned sinusoid and square-wave stimuli might be interesting to disambiguate orientation effects (alignment) and stronger binding of luminance and colour; although a filling-in effect might take place.

Stimuli for further studies Some extensions were already suggested in this chapter (e.g. testing the suprathreshold contrast perception effect with different viewing conditions, binocular or dichoptic viewing, and using rendered objects or objects that are more realistic). Here, I suggest some other stimuli that would bring new information to luminance-chromatic interactions.

It would be interesting to see the effect of differential SFs between target and mask.

[11] Only the fundamental frequency is conserved.

[12] As discussed in Chapter 6, perceived contrast of a square-wave vs. sine-wave is increased by a constant ratio.

Testing suprathreshold contrast with a difference between test and mask SF would extend the range of conditions tested. Furthermore, comparing between variations in orientation and SF, of luminance target and colour mask would be of interest in the detection task. This would compare orientation channels with SF channel interactions and tell us about the general connectivity of spatial processing. I would predict that the highest suppression possibly occurs at similar SFs, as iso-orientation has often the greater suppressive effect (see Section 3.3.3, p. 36).

In the case of the colour-shading experiment, in terms of heuristics, SF should not play a role; so this effect should be similar with different SFs. However, if differential SF processing does exist, it might just be resulting from low-level interactions between SF channels without a specific computational goal of extracting reflectance vs. shape as I hypothesized. The dependency of the background colour would also be interesting to test (see HK effect, Section 9.2.2). I would expect a possibly stronger effect with coloured background.

In order to learn more about the underlying physiology of the channel interactions, I would suggest performing frequency tagging steady-state visually evoked potential (ssVEP) EEG study. In this experiment, I would test chromatic and luminance at several relative orientations and SFs (Burr & Morrone, 1987; Candy, Skoczenski, & Norcia, 2001; Norcia, 2004; Regan & Regan, 1987). The frequency interactions in the frequency domain would be of particular interest (compare to luminance interactions with orthogonal stimuli Candy et al., 2001). A sweep parameter ssVEP with increased contrast on the mask (sweep parameter) would provide direct evidence of connectivity performing gain control (Norcia, 2004) between luminance and chromatic components. Furthermore, I would look for variations in these gain controls between participants as was found in this thesis, these would be the physiological correlates of the psychophysical results shown in this thesis. As this methodology does not necessarily require a task, it could also demonstrate conclusively that the results are not due to strategy.

9.3 General conclusion

In this thesis, I tested several paradigms in which chromatic and luminance information might interact in the processing of spatial information (Section 9.1.1). The data showed that chromatic information did interact with luminance processing at the suprathreshold level; this interaction was found both for shape-from-shading (colour-shading effect) and for suprathreshold contrast perception. There was a wide repertoire of behaviours between participants, which led me to ask about the possible role of these interactions but also how they come to be. I also presented data in isoluminance with different paradigms (flicker vs. motion).

The findings regarding chromo-luminance interactions were previously interpreted with the heuristic hypothesis of processing reflectance changes and shape changes. It is a difficult task to assess if a particular effect is due to a processing rule or if it might be emerging

from a different computation. I have argued that the heuristics interpretation should be taken with caution, especially in its genetic formulation (Ramachandran, 1991). However, I have argued that these interactions might emerge from the visual environment statistics and could be more flexible. This hypothesis should be tested with further experiments on long and medium term perceptual learning.

The data at the detection threshold did reveal very few interactions between luminance and colour. However, it was tested only in one direction, colour effects on luminance, and the reverse direction might not be so (see Section 3.4.4, p. 51). This specific combination was studied because, as mentioned in this chapter, luminance information tends to be integrated faster for objects and spatial processing. However, isoluminant information can allow similar complex processing (i.e. objects). Spatial processing between different channels is a subject of interest and there is still an open question about how several channels (SF, orientation within luminance red-green and blue-yellow) are combined together to form a unified suprathreshold perception. The combination of all these features requires not only the study of each of these channels on their own but also their interactions and interactions in the context in which they are used: the visual understanding of the environment. Studies such as the ones in this thesis bring some understanding on how our seemingly rich visual experience is processed by interacting channels at different levels of perception (detection, suprathreshold, and higher level tasks such as curvature analysis) and, interestingly, it outlines the individuality of different observers' visual systems.

References

Adelson, E. H. (1993). Perceptual organization and the judgment of brightness. *Science*, *262*, 2042–2044.

Adelson, E. H., & Pentland, A. P. (1996). Perception as bayesian inference. In D. Knill & W. Richards (Eds.), (pp. 409–423). New York: Cambridge University Press.

Albrecht, D. G. (1995). Visual cortex neurons in monkey and cat: Effect of contrast on the spatial and temporal phase transfer functions. *Visual Neuroscience*, *12*, 1191–1210.

Albrecht, D. G., & Hamilton, D. B. (1982). Striate cortex of monkey and cat: Contrast response function. *Journal of Neurophysiology*, *48*(1), 217–237.

Anderson, J. S., Carandini, M., & Ferster, D. (2000). Orientation tuning of input conductance, excitation, and inhibition in cat primary visual cortex. *Journal of Neurophysiology*, *84*(2), 909–926.

Angelucci, A., & Bullier, J. (2003). Reaching beyond the classical receptive field of v1 neurons: horizontal or feedback axons? *Journal of Physiology Paris*, *97*(2-3), 141–154.

Anstis, S., & Cavanagh, P. (1983). A minimum motion technique for judging equiluminance. In J. Mollon & L. Sharpe (Eds.), *Colour vision: Physiology and psychophysics* (pp. 165–166). London:Academic press.

Barrow, H., & Tenenbaum, J. (1978). Computer vision systems. In A. Hanson & E. Riseman (Eds.), *Computer vision systems* (pp. 3–26). New York: Academic Press.

Bauman, L. A., & Bonds, A. B. (1991). Inhibitory refinement of spatial frequency selectivity in single cells of the cat striate cortex. *Vision Research*, *31*(6), 933–44.

Baylor, D. A. (1987). Photoreceptor signals and vision: Proctor lecture. *Investiguative Ophtalmology and Visual Science*, *28*(1), 34–49.

Baylor, D. A., Hodgkin, A. L., & Lam, T. D. (1974). The electrical response of turtle cones to flashes and steps of light. *The Journal of Physiology*, *242*, 685–727.

Benardete, E. A., & Kaplan, E. (1999). The dynamics of primate m retinal ganglion cells. *Visual Neuroscience*, *16*, 355–368.

Billock, V. A. (1995). Cortical simple cells can extract achromatic information from the multiplexed chromatic and achromatic signals in the parvocellular pathway. *Vision Research*, *35*(16), 2359–2369.

Billock, V. A., & Tsou, B. H. (2004). A role for cortical crosstalk in the binding problem: stimulus-driven correlations that link color, form, and motion. *Journal of Cognitive*

Neuroscience, *16*(4), 1036–1048.

Bird, C., Henning, G., & Wichmann, F. (2002). Contrast discrimination with sinusoidal gratings of different spatial frequency. *Journal of the Optical Society of America A - Optics, Image Science and Vision*, *19*(7), 1267–1273.

Blakemore, C., & Campbell, F. W. (1969). On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images. *Journal of Physiology*, *203*(1), 237–260.

Blakemore, C., Muncey, J. P. J., & Ridley, R. M. (1973). Stimulus specificity in the human visual system. *Vision Research*, *13*(10), 1915–1931.

Blaser, E., Sperling, G., & Lu, Z.-L. (1997). Measuring the spatial resolution of visual attention. *Investiguative Ophtalmology and Visual Science, ARVO Supplement*, *38*(4).

Bloj, M., Kersten, D., & Hurlbert, A. (1999). Perception of three-dimensional shape influences colour perception through mutual illumination. *Nature*, *402*(6764), 877–879.

Booker, R. L. (1981). Luminancebrightness comparisons of separated circular stimuli. *Journal of the Optical Society of America*, *71*(2), 139–144.

Boycott, B. B., & Wässle, H. (1991). Morphological classification of bipolar cells of the primate retina. *European Journal of Neuroscience*, *3*(11), 1069–1088.

Bradley, A., Switkes, E., & Valois, K. D. (1988). Orientation and spatial frequency selectivity of adaptation to color and luminance gratings. *Vision Research*, *28*(7), 841–856.

Brainard, D. H. (1996). Cone contrast and opponent modulation color spaces. In P. K. Kaiser & R. M. Boynton (Eds.), *Human color vision* (Second Edition ed., pp. 563–579). Optical Society of America.

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

Brainard, D. H., Pelli, D. G., & Robson, T. (2002). Display characterization. In H. J (Ed.), *Encyclopedia of imaging science and technology* (pp. 172–188). Wiley.

Brainard, D. H., Roorda, A., Yamauchi, Y., Calderone, J. B., Metha, A., Neitz, M., ... Jacobs, G. H. (2000). Functional consequences of the relative numbers of L and M cones. *Journal of the Optical Society of America A-optics Image Science and Vision*, *17*(3), 607-614.

Brandon, C., & Lam, D. M. (1983). L-glutamic acid: A neurotransmitter candidate for cone photoreceptors in human and rat retinas. *Proceedings of the National Academy of Sciences*, *80*(16), 5117–5121.

Brannan, J. R., & Bodis-Wollner, I. (1991). Evidence for two systems mediating perceived contrast. *Visual Neuroscience*, *6*, 587–592.

Bülthoff, H. H., & Mallot, H. A. (1988). Integration of depth modules: Stereo and shading. *Journal of the Optical Society of America A*, *5*, 1749–1758.

Burns, S. A., Smith, V. C., Pokorny, J., & Elsner, A. E. (1982). Brightness of equal-luminance lights. *Journal of the Optical Society of America*, *72*(9), 1225–1231.

Burr, D. C., & Morrone, C. (1993). Impulse-response functions for chromatic and achromatic stimuli. *Journal of the Optical Society of America*, *10*(8), 1706–1713.

Burr, D. C., & Morrone, M. C. (1987). Inhibitory interactions in the human vision system revealed in pattern-evoked potentials. *Journal of Physiology*, *389*, 1–21.

Campbell, F. W., & Robson, J. G. (1968). Application of fourier analysis to the visibility of gratings. *Journal of Physiology*, *197*(3), 551–566.

Candy, T. R., Skoczenski, A. M., & Norcia, A. M. (2001). Normalization models applied to orientation masking in the human infant. *Journal of Neuroscience*, *21*(12), 4530–4541.

Cannon, M. W., & Fullenkamp, S. C. (1991a). Spatial interactions in apparent contrast: inhibitory effects among grating patterns of different spatial frequencies, spatial positions and orientations. *Vision Research*, *31*(11), 1985–1998.

Cannon, M. W., & Fullenkamp, S. C. (1991b). A transducer model for contrast perception. *Vision Research*, *31*(6), 983–998.

Cannon, M. W., & Fullenkamp, S. C. (1993). Spatial interactions in apparent contrast: Individual differences in enhancement and suppression effects. *Vision Research*, *33*(12), 1685–1695.

Carandini, M., & Heeger, D. J. (2012). Normalization as a canonical neural computation. *Nature Reviews Neuroscience*, *13*, 51–62.

Carandini, M., Heeger, D. J., & Movshon, J. A. (1997). Linearity and normalization in simple cells of the macaque primary visual cortex. *The Journal of Neuroscience*, *17*(21), 8621–8644.

Cass, J., & Alais, D. (2006). Evidence for two interacting temporal channels in human visual processing. *Vision Reserach*, *46*(18), 2859–2868.

Cass, J., Clifford, C. W. G., Alais, D., & Spehar, B. (2009). Temporal structure of chromatic channels revealed through masking. *Journal of Vision*, *9*(5).

Cavanaugh, J. R., Bair, W., & Movshon, J. A. (2002). Selectivity and spatial distribution of signals from the receptive field surround in macaque V1 neurons. *Journal of Neurophysiology*, *88*(5), 2547–2556.

Chan, T. L., Martin, P. R., Clunas, N., & Grünert, U. (2001). Bipolar cell diversity in the primate retina: Morphologic and immunocytochemical analysis of a new world monkey, the marmoset callithrix jacchus. *The Journal of Comparative Neurology*, *437*, 219–239.

Chatterjee, S., & Callaway, E. M. (2003). Parallel colour-opponent pathways to primary visual cortex. *Nature*, *426*(6967), 668–671.

Chen, C. C., Foley, J. M., & Brainard, D. H. (2000a). Detection of chromoluminance patterns on chromoluminance pedestals I: Threshold measurements. *Vision Research*, *40*(7), 773–788.

Chen, C. C., Foley, J. M., & Brainard, D. H. (2000b). Detection of chromoluminance patterns on chromoluminance pedestals II: Model. *Vision Research*, *40*(7), 789–803.

CIE. (1932). *Commission internationale de l´Éclairage proceedings, 1931.* Cambridge: Cambridge University Press.

Clery, S., Bloj, M., & Harris, J. M. (2013). Interactions between luminance and color signals: Effects on shape. *Journal of Vision*, *13*(5).

Clifford, C. W. G., Spehar, B., Solomon, S. G., Martin, P. R., & Zaidi, Q. (2003). Interactions between color and luminance in the perception of orientation. *Journal of Vision*, *3*(2).

Cole, G. R., & Hine, T. (1992). Computation of cone contrasts for color-vision research. *Behavior Research Methods Instruments & Computers*, *24*(1), 22–27.

Cole, G. R., Hine, T., & McIlhagga, W. H. (1993). Detection mechanisms in L-cone,M-cone,and S-cone contrast space. *Journal of the Optical Society of America A - Optics, Image Science and Vision*, *10*(1), 38–51.

Cole, G. R., Stromeyer, C. F., & Kronauer, R. E. (1990). Visual interactions with luminance and chromatic stimuli. *Journal of the Optical Society of America A - Optics, Image Science and Vision*, *7*(1), 128–140.

Copenhagen, D. R., & Jahr, C. E. (1989). Release of endogenous excitatory amino acids from turtle photoreceptors. *Nature*, *341*(6242), 536–539.

Crook, J. M., & Eysel, U. T. (1992). GABA-induced inactivation of functionally characterized sites in cat visual cortex (Area 18): Effects on orientation tuning. *The Journal of Neuroscience*, *12*(5), 1816–1825.

Curran, W., & Johnston, A. (1996). The effect of illuminant position on perceived curvature. *Vision Research*, *36*(10), 1399–1410.

Dacey, D. M. (1993). Morphology of a small-field bistratified ganglion cell type in the macaque and human retina. *Visual Neuroscience*, *10*(6), 1081–1098.

Dacey, D. M. (2000). Parallel pathways for spectral coding in primate retina. *Annual Review of Neuroscience*, *23*, 743–775.

Dacey, D. M., & Lee, B. B. (1994). The 'blue-on' opponent pathway in primate retina originates from a distinct bistratified ganglion cell type. *Nature*, *367*(6465), 731–735.

Dacey, D. M., & Packer, O. S. (2003). Colour coding in the primate retina: Diverse cell types and cone-specific circuitry. *Current Opinion in Neurobiology*, *13*, 421–427.

Dacey, D. M., Packer, O. S., Diller, L., Brainard, D. H., & Lee, B. B. (2000). Center surround receptive field structure of cone bipolar cells in primate retina. *Vision Research*, *40*, 1801–1811.

Dacey, D. M., Peterson, B., Robinson, F., & Gamlin, P. (2003). Fireworks in the primate retina: in vitro photodynamics reveals diverse LGN-projecting ganglion cell types. *Neuron*, *37*, 15–27.

Dean, A. F. (1981). The relationship between response amplitude and contrast for cat striate cortical neurones. *Journal of Physiology*, *318*, 413–427.

DeAngelis, G. C., Freeman, R. D., & Ohzawa, I. (1994). Length and width tuning of neurons in the cat's primary visual cortex. *Journal of Neurophysiology*, *71*(1), 347–

374.

DeAngelis, G. C., Robson, J. G., Ohzawa, I., & Freeman, R. D. (1992). Organization of suppression in receptive fields of neurons in cat visual cortex. *Journal of Neurophysiology*, *68*(1), 144–163.

Derrington, A. M., Krauskopf, J., & Lennie, P. (1984). Chromatic mechanisms in lateral geniculate-nucleus of macaque. *Journal of Physiology*, *357*, 241–265.

Derrington, A. M., & Lennie, P. (1984). Spatial and temporal contrast sensitivities of neurones in lateral geniculate nucleus of macaque. *Journal of Physiology*, *357*, 219–240.

DeValois, K. K., & Switkes, E. (1983). Simultaneous masking interactions between chromatic and luminance gratings. *Journal of the Optical Society of America*, *73*(1), 11–18.

de Vries, H. L. (1948). The fundamental response curves of normal and abnormal dichromatic and trichromatic eyes. *Physica*, *14(6)*, 367–380.

Diller, L., Packer, O. S., Verweij, J., McMahon, M. J., Williams, D. R., & Dacey, D. M. (2004). L and M cone contributions to the midget and parasol ganglion cell receptive fields of macaque monkey retina. *The Journal of Neuroscience*, *24*(5), 1079–1088.

Dobkins, K. R., Anderson, C. M., & Kelly, J. (2001). Development of psychophysically-derived detection contours in l- and m-cone contrast space. *Vision Research*, *41*(14), 1791–1807.

Dobkins, K. R., Thiele, A., & Albright, T. D. (2000). Comparison of red-green equiluminance points in humans and macaques: evidence for different L : M cone ratios between species. *Journal of the Optical Society of America*, *17*(3), 545–556.

Dovencioglu, D. N., Schofield, A. J., & Welchman, A. E. (2011). Learning to see second order information in shading patterns. *Journal of Vision*, *11*(11), 1031.

Dovencioglu, D. N., Welchman, A. E., & Schofield, A. J. (2013). Perceptual learning of second order cues for layer decomposition. *Vision Research*, *77*, 1–9.

Dowling, J. E. (1968). Synaptic organization of the frog retina: An electron microscopic analysis comparing the retinas of frogs and primates. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *170*(1019), 205–228.

Einstein, A. (1905). Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt. *Annalen der Physik*, *17*(6), 132-148.

Ejima, Y., & Takahashi, S. (1985). Apparent contrast of sinusoidal grating in the simultaneous presence of peripheral gratings. *Vision Research*, *25*(9), 1223–1232.

Elsner, A. E., Pokorny, J., & Burns, S. A. (1986). Chromaticity discrimination: Effects of luminance contrast and spatial frequency. *Journal of the Optical Society of America*, *3*(7), 916–920.

Eskew, R. T., Stromeyer, C. F., Picotte, C. J., & Kronauer, R. E. (1991). Detection uncertainty and the facilitation of chromatic detection by luminance contours. *Journal of the Optical Society of America*, *8*(2), 394–403.

Famiglietti, E. V., & Kolb, H. (1976). Structural basis for ON- and OFF-center responses

in retinal ganglion cells. *Science*, *194*(4261), 193–195.

Favreau, O. E., & Cavanagh, P. (1981). Color and luminance: Independent frequency shifts. *Science*, *212*, 831–832.

Favreau, O. E., & Cavanagh, P. (1983). Interocular transfer of a chromatic frequency shift. *Vision Research*, *23*, 951–957.

Field, G. D., Gauthier, J. L., Sher, A., Greschner, M., Machado, T. A., Jepson, L. H., ... Chichilnisky, E. J. (2010). Functional connectivity in the retina at the resolution of photoreceptors. *Nature*, *467*, 673–677.

Flanagan, P., Cavanagh, P., & Favreau, O. E. (1990). Independent orientation-selective mechanisms for the cardinal directions of colour space. *Vision Research*, *30*, 769–778.

Foley, J. M. (1994). Human luminance pattern-vision mechanisms: Masking experiments require a new model. *Journal of the Optical Society of America A*, *11*, 1710–1719.

Foley, J. M., & Legge, G. E. (1981). Contrast detection and near-threshold discrimination in human vision. *Vision Research*, *21*(7), 1041–1053.

Friedman, H. S., Zhou, H., & von der Heydt, R. (2003). The coding of uniform colour figures in monkey visual cortex. *The Journal of Physiology*, *548*, 593–613.

Georgeson, M. A., & Meese, T. S. (1997). Perception of stationary plaids: The role of spatial filters in edge analysis. *Vision Research*, *37*(23), 3255–3271.

Georgeson, M. A., & Shackleton, T. M. (1994). Perceived contrast of gratings and plaids: Nonlinear summation across oriented filters. *Vision Research*, *34*(8), 1061–1075.

Georgeson, M. A., & Sullivan, G. D. (1975). Contrast constancy: Deblurring in human vision by spatial frequency channels. *Journal of Physiology*, *252*(3), 627–656.

Gheiratmand, M., Meese, T. S., & Mullen, K. T. (2013). Blobs versus bars: Psychophysical evidence supports two types of orientation response in human color vision. *Journal of Vision*, *13*.

Gheiratmand, M., & Mullen, K. T. (2010). Cortical color-luminance contrast interactions revealed by dichoptic masking. *Journal of Vision*, *10*(15), 55.

Gilchrist, A. L. (1977). Perceived lightness depends on perceived spatial arrangement. *Science*, *195*(4274), 185–187.

Gilchrist, A. L., Delman, S., & Jacobsen, A. (1983). The classification and integration of edges as critical to the perception of reflectance and illumination. *Perception & Psychophysics*, *33*(5), 425–436.

Ginsburg, A. P., Cannon, M. W., & Nelson, M. A. (1980). Suprathreshold processing of complex visual-stimuli: Evidence for linearity in contrast perception. *Science*, *208*(4444), 619–621.

Giulianini, F., & Eskew, R. T. (1998). Chromatic masking in the (ΔL/L, ΔM/M) plane of cone-contrast space reveals only two detection mechanisms. *Vision Research*, *38*(24), 3913–3926.

Gowdy, P., Stromeyer, C., & Kronauer, R. (1999). Facilitation between the luminance and red-green detection mechanisms: enhancing contrast differences across edges.

Vision Research, *39*(24), 4098–4112.

Granzier, J. J. M., & Gegenfurtner, K. R. (2012). Effects of memory colour on colour constancy for unknown coloured objects. *i-Perception*, *3*, 190–215.

Green, D. M., & Luce, R. D. (1974). Variability of magnitude estimates: a timing theory analysis. *Perception & Psychophysics*, *15*(2), 291–300.

Guild, J. (1931). The colorimetric properties of the spectrum. *Philosophical Transactions of the Royal Society of London*, *A230*, 149–187.

Gunther, K. L., & Dobkins, K. R. (2002). Individual differences in chromatic (red/green) contrast sensitivity are constrained by the relative number of L- versus M-cones in the eye. *Vision Research*, *42*(11), 1367–1378.

Gur, M., & Akri, V. (1992). Isoluminant stimuli may not expose the full contribution of color to visual functioning: spatial contrast sensitivity measurements indicate interaction between color and luminance processing. *Vision Research*, *32*(7), 1253–1262.

Guth, S. L., Donley, N. J., & Marrocco, R. T. (1969). On luminance additivity and realted topics. *Vision Research*, *9*, 537–575.

Hamerly, J. R., Quick, R. F., & Reichert, T. A. (1977). Study of grating contrast judgment. *Vision Research*, *17*(2), 201–207.

Hansen, T., & Gegenfurtner, K. R. (2009). Independence of color and luminance edges in natural scenes. *Visual Neuroscience*, *26*(1), 35–49.

Harding, G., Harris, J. M., & Bloj, M. (2012). Learning to use illumination gradients as an unambiguous cue to three dimensional shape. *PLoS ONE*, *7*(4).

Hecht, S. (1937). Rods, cones and the chemical basis of vision. *Physiological Reviews*, *17*, 239–290.

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

Hering, E. (1964). *Outlines of a theory of the light sense* (L. M. Hurvitch & D. Jameson, Trans.). Harvard, Cambdridge, Mass.

Hilz, R. L., & Cavonius, C. R. (1970). Wavelength discrimination measured with square-wave gratings. *Journal of the Optical Society of America*, *60*(2), 273–277.

Hilz, R. L., Huppmann, G., & Cavonius, C. R. (1974). Influence of luminance contrast on hue discrimination. *Journal of the Optical Society of America*, *64*(6), 763–766.

Horn, B. K. P. (1974). Determining lightness from an image. *Computer Graphics and Image Processing*, *3*, 277–299.

Horn, B. K. P. (1975). Obtaining shape from shading information. In P. H. Winston (Ed.), *Psychology of computer vision* (pp. 115–155). New York: McGraw-Hill.

Horn, B. K. P., & Brooks, M. J. (Eds.). (1989). *Shape from shading.* Cambridge, Mass.: The MIT Press.

Hou, C., Pettet, M. W., Vildavski, V. Y., & Norcia, A. M. (2006). Neural correlates of shape-from-shading. *Vision Research*, *46*(6–7), 1080–1090.

Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional

architecture in the cat's visual cortex. *Journal of Neurophysiology*, *160*, 106–154.

Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *Journal of Physiology*, *195*, 215–243.

Hubel, D. H., & Wiesel, T. N. (1977). Ferrier lecture: Functional architecture of macaque monkey visual cortex. *Proceedings of the Royal Society of London B*, *198*(1130), 1–59.

Ichida, J. M., Schwabe, L., Bressloff, P. C., & Angelucci, A. (2007). Response facilitation from the "suppressive" receptive field surround of macaque V1 neurons. *Journal of Neurophysiology*, *98*, 2168–2181.

Ingling, C. R., & Martinez-Uriegas, E. (1983). The relationship between spectral sensitivity and spatial sensitivity for the primate r-g X-channel. *Vision Research*, *23*(12), 1495–1500.

Jackman, S. L., Babai, N., Chambers, J. J., Thoreson, W. B., & Kramer, R. H. (2011). A positive feedback synapse from retinal horizontal cells to cone photoreceptors. *PLoS Biol*, *9*(5).

Jagla, W. M., Jägle, H., Hayashi, T., Sharpe, L. T., & Deeb, S. S. (2002). The molecular basis of dichromatic color vision in males with multiple red and green visual pigment genes. *Human Molecular Genetics*, *11*(1), 23–32.

Jägle, H., de Luca, E., Serey, L., Bach, M., & Sharpe, L. T. (2006). Visual acuity and X-linked color blindness. *Graefe's Archive for Clinical and Experimental Ophthalmology*, *244*, 447–453.

Johnson, E. N., Hawken, M. J., & Shapley, R. (2001). The spatial transformation of color in the primary visual cortex of the macaque monkey. *Nature Neuroscience*, *4*(4), 409–416.

Johnson, E. N., Hawken, M. J., & Shapley, R. (2004). Cone inputs in macaque primary visual cortex. *Journal of Neurophysiology*, *91*, 2501–2514.

Jones, H. E., Wang, W., & Sillito, A. M. (2002). Spatial organization and magnitude of orientation contrast interactions in primate V1. *Journal of Neurophysiology*, *88*(5), 2796–2808.

Jordan, G., Deeb, S. S., Bosten, J. M., & Mollon, J. D. (2010). The dimensionality of color vision in carriers of anomalous trichromacy. *Journal of Vision*, *10*(8).

Judd, D. B. (1951). Report of u.s. secretariat committee on colorimetry and artificial daylight. *Proceedings of the Twelfth Session of the CIE*, 1–11.

Judd, D. B., MacAdam, D. L., & Wyszecki, G. (1964). Spectral distribution of typical daylight as a function of correlated color temperature. *Journal of the Optical Society of America*, *54*(8), 1032–1039.

Julesz, B. (1971). *Foundations of cyclopean perception.* Chicago University Press.

Kaplan, E., & Shapley, R. M. (1986). The primate retina contains two types of ganglion cells, with high and low contrast sensitivity. *Proceedings of the National Academy of Sciences*, *83*, 2755–2757.

Kelly, D. H. (1983). Spatiotemporal variation of chromatic and achromatic contrast

thresholds. *Journal of the Optical Society of America*, *73*(6), 742–750.

Kerrigan, I. S., & Adams, W. J. (2013). Highlights, disparity, and perceived gloss with convex and concave surfaces. *Journal of Vision*, *13*.

Kim, Y. J., Gheiratmand, M., & Mullen, K. T. (2013). Cross-orientation masking in human color vision: Application of a two-stage model to assess dichoptic and monocular sources of suppression. *Journal of Vision*, *13*(6).

Kimura, R., & Ohzawa, I. (2009). Time course of cross-orientation suppression in the early visual cortex. *Journal of Neurophysiology*, *101*, 1463–1479.

Kingdom, F. A. A. (1999). Old wine in new bottles? some thoughts on logvinenkos lightness induction revisited. *Perception*, *28*, 929–934.

Kingdom, F. A. A. (2003). Color brings relief to human vision. *Nature Neuroscience*, *6*(6), 641–644.

Kingdom, F. A. A. (2008). Perceiving light versus material. *Vision Research*, *48*(20), 2090–2105.

Kingdom, F. A. A., Bell, J., Gheorghiu, E., & Malkoc, G. (2010). Chromatic variations suppress suprathreshold brightness variations. *Journal of Vision*, *10*(13), 1–13.

Kingdom, F. A. A., Rangwala, S., & Hammamji, K. (2005). Chromatic properties of the colour-shading effect. *Vision Research*, *45*(11), 1425–1437.

Kingdom, F. A. A., & Simmons, D. R. (1998). The missing-fundamental illusion at isoluminance. *Perception*, *27*(12), 1451–1460.

Kingdom, F. A. A., Simmons, D. R., & Rainville, S. (1999). On the apparent collapse of stereopsis in random-dot-stereograms at isoluminance. *Vision Research*, *39*(12), 2127–2141.

Kingdom, F. A. A., Wong, K., Yoonessi, A., & Malkoc, G. (2006). Colour contrast influences perceived shape in combined shading and texture patterns. *Spatial Vision*, *19*(2-4).

Kleffner, D. A., & Ramachandran, V. S. (1992). On the perception of shape from shading. *Perception & Psychophysics*, *52*(1), 18–36.

Klug, K., Herr, S., Ngo, I. T., Sterling, P., & Schien, S. (2003). Macaque retina contains an S-cone OFF midget pathway. *The Journal of Neuroscience*, *23*(30), 9881–9887.

Knill, D. C., & Kersten, D. (1991). Apparent surface curvature affects lightness perception. *Nature*, *351*, 229–230.

Kolb, H. (1970). Organization of the outer plexiform layer of the primate retina: electron microscopy of golgi-impregnated cells. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *255*, 1–14.

Kolb, H. (1997). Amacrine cells of the mammalian retina: Neurocircuitry and functional roles. *Eye*, *11*, 904–923.

Kolb, H., & Dekorver, L. (1991). Midget ganglion cells of the parafovea of the human retina: a study by electron microscopy and serial section reconstructions. *The Journal of Comparative Neurology*, *303*(4), 617–636.

Kolb, H., Fernandez, E., Schouten, J., Ahnelt, P., Llnberg, K. A., & Fisher, S. K. (1994).

Are there three types of horizontal cell in the human retina? *The Journal of Comparative Neurology*, *343*, 370–386.

Kontsevich, L., & Tyler, C. W. (1999). Bayesian adaptive estimation of psychometric slope and threshold. *Vision Research*, *39*, 2729–2737.

Kraft, T. W., Schneeweis, D. M., & Schnapf, J. L. (1993). Visual transduction in human rod photoreceptors. *Journal of Physiology*, *464*, 747–765.

Krauskopf, J., Williams, D. R., & Heeley, D. W. (1982). Cardinal directions of color space. *Vision Research*, *22*(9), 1123–1131.

Kremers, J., Scholl, H., Knau, H., Berendschot, T., Usui, T., & Sharpe, L. (2000). L/M cone ratios in human trichromats assessed by psychophysics,electroretinography,and retinal densitometry. *Journal of the Optical Society of America A - Optics, Image Science and Vision*, *17*(3), 517–526.

Kremers, J., Usui, T., Scholl, H. P. N., & Sharpe, L. T. (1999). Cone signal contributions to electrograms in dichromats and trichromats. *Investigative Ophthalmology & Visual Science*, *40*(5), 920–930.

Land, E. H. (1977). The retinex theory of color vision. *Scientific American*, *237*(6), 108–128.

Langer, M. S., & Bülthoff, H. H. (2000). Depth discrimination from shading under diffuse lighting. *Perception*, *29*(6), 649–660.

Langer, M. S., & Bülthoff, H. H. (2001). A prior for global convexity in local shape-from-shading. *Perception*, *30*(4), 403–410.

Lee, B. B. (2008). Neural models and physiological reality. *Visual Neuroscience*, *25*, 231–240.

Lee, B. B. (2011). Visual pathways and psychophysical channels in the primate. *Journal of Physiology*, *589*, 41–47.

Lee, B. B., Martin, P. R., & Grünert, U. (2010). Retinal connectivity and primate vision. *Progress in Retinal and Eye Research*, *29*, 622–639.

Lee, B. B., Martin, P. R., & Valberg, A. (1988). The physiological basis of heterochromatic flicker photometry demonstrated in the ganglion cells of the macaque retina. *Journal of Physiology*, *404*, 223–243.

Lee, B. B., Sun, H., & Valberg, A. (2011). Segregation of chromatic and luminance signals using a novel grating stimulus. *The Journal of Physiology*, *589*(1), 59–73.

Legge, G. E., & Foley, J. M. (1980). Contrast masking in human-vision. *Journal of the Optical Society of America*, *70*(12), 1458–1471.

Lennie, P. (2003). The physiology of color vision. In S. K. Shevell (Ed.), *The science of color* (2nd ed., pp. 217–246). Elsevier.

Lennie, P., Krauskopf, J., & Sclar, G. (1990). Chromatic mechanisms in striate cortex of macaque. *Journal of Neuroscience*, *10*(2), 649–669.

Lennie, P., Pokorny, J., & Smith, V. C. (1993). Luminance. *Journal of the Optical Society of America A*, *10*(6), 1283–1293.

Leventhal, A. G., Rodieck, R. W., & Dreher, B. (1981). Retinal ganglion cell classes in

the old world monkey: Morphology and central projections. *Science*, *213*(4512), 1139–1142.

Leventhal, A. G., Thompson, K. G., Liu, D., Zhou, Y., & Ault, S. J. (1995). Concomitant sensitivity to orientation, direction, and color of cells in layers 2, 3, and 4 of monkey striate cortex. *The Journal of Neuroscience*, *15*(3), 1808–1818.

Levitt, J. B., & Lund, J. S. (1997). Contrast dependence of contextual effects in primate visual cortex. *Nature*, *387*(6628), 73–76.

Livingstone, M. S., & Hubel, D. H. (1984). Anatomy and physiology of a color system in the primate visual cortex. *Journal of Neuroscience*, *4*(1), 309–356.

Livingstone, M. S., & Hubel, D. H. (1987). Psychophysical evidence for separate channels for the perception of form, color, movement, and depth. *Journal of Neuroscience*, *7*, 3416–3468.

Livingstone, M. S., & Hubel, D. H. (1988). Segregation of form, color, movement, and depth-anatomy, physiology, and perception. *Science*, *240*(4853), 740–749.

Logvinenko, A. D. (1999). Lightness induction revisited. *Perception*, *28*, 803–816.

Losada, M. A., & Mullen, K. T. (1994). The spatial tuning of chromatic mechanisms identified by simultaneous masking. *Vision Research*, *34*(3), 331–341.

Lotto, R. B., & Purves, D. (1999). The effects of color on brightness. *Nature Neuroscience*, *2*(11), 1010–1014.

Lovell, P. G., Bloj, M., & Harris, J. M. (2012). Optimal integration of shading and binocular disparity for depth perception. *Journal of Vision*, *12*(1).

Lu, C., & Fender, D. H. (1972). Interaction of color and luminance in stereoscopic vision. *Investigative Ophtamology*, *11*(6), 482–490.

Lu, Z.-L., Lesmes, L. A., & Sperling, G. (1999). The mechanism of isoluminant chromatic motion perception. *Proceedings of the National Academy of Sciences*, *96*, 8289–8294.

Lund, J. S. (1988). Anatomical organization of macaque monkey striate visual cortex. *Annual Review of Neuroscience*, *11*, 253–288.

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

Marr, D., & Hildreth, E. (1980). Theory of edge detection. *Proceedings of the Royal Society of London B*, *207*(187–217).

Martinovic, J., Mordal, J., & Wuerger, S. M. (2011). Event-related potentials reveal an early advantage for luminance contours in the processing of objects. *Journal of Vision*, *11*(7).

Marčelja, S. (1980). Mathematical description of the responses of simple cortical cells. *Journal of the Optical Society of America*, *70*(11), 1297–1300.

McCollough, C. (1965). Color adaptation of edge detectors in the human visual system. *Science*, *149*, 1115–1116.

McIlhagga, W. H., & Mullen, K. T. (1996). Contour integration with colour and luminance contrast. *Vision Research*, *36*, 1995–200.

McKeefry, D., Parry, N., & Murray, I. (2003). Simple reaction times in color space:

The influence of chromaticity, contrast, and cone opponency. *Investigative Ophthalmology & Visual Science*, *44*(5), 2267-2276. (Annual Meeting of the Association-for-Research-in-Vision-and-Ophthalmology, FT LAUDERDALE, FLORIDA, MAY 05-10, 2002)

McNaughton, P. A. (1990). The light response of photoreceptors. In C. Blakemore (Ed.), *Vision: coding and efficiency.*

Medina, J. M., & Diaz, J. A. (2006). Postreceptoral chromatic-adaptation mechanisms in the red-green and blue-yellow systems using simple reaction times. *Journal of the Optical Society of America A - Optics, Image Science and Vision*, *23*(5), 993–1007.

Medina, J. M., & Mullen, K. T. (2009). Cross-orientation masking in human color vision. *Journal of Vision*, *9*(3), 1–16.

Meese, T. S., & Baker, D. H. (2011). Contrast summation across eyes and space is revealed along the entire dipper function by a swiss cheese stimulus. *Journal of Vision*, *11*(1).

Meese, T. S., Challinor, K. L., Summers, R. J., & Baker, D. H. (2009). Suppression pathways saturate with contrast for parallel surrounds but not for superimposed cross-oriented masks. *Vision Research*, *49*(24), 2927–2935.

Meese, T. S., & Hess, R. F. (2004). Low spatial frequencies are suppressively masked across spatial scale, orientation, field position, and eye of origin. *Journal of Vision*, *4*, 843–859.

Meese, T. S., & Holmes, D. J. (2007). Spatial and temporal dependencies of cross-orientation suppression in human vision. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *274*(1606), 127–136.

Meese, T. S., & Holmes, D. J. (2010). Orientation masking and cross-orientation suppression (xos): Implications for estimates of filter bandwidth. *Journal of Vision*, *10*(12), 1–20.

Meese, T. S., & Summers, R. J. (2007). Area summation in human vision at and above detection threshold. *Proceedings of the Royal Society of London B*, *274*(1627), 2891–2900.

Meese, T. S., & Summers, R. J. (2012). Theory and data for area summation of contrast with and without uncertainty: Evidence for a noisy energy model. *Journal of Vision*, *12*(11).

Meese, T. S., Summers, R. J., Holmes, D. J., & Wallis, S. A. (2007). Contextual modulation involves suppression and facilitation from the center and the surround. *Journal of Vision*, *7*(4).

Metha, A. B., & Mullen, K. T. (1997). Redgreen and achromatic temporal filters: a ratio model predicts contrast-dependent speed perception. *Journal of the Optical Society of America*, *14*(5), 984–996.

Michaelides, M., Johnson, S., Simunovic, M. P., Bradshaw, K., Holder, G., Mollon, J. D., ... Hunt, D. M. (2005). Blue cone monochromatism: a phenotype and genotype assessment with evidence of progressive loss of cone function in older individuals. *Eye*, *19*, 2–10.

Michelson, A. A. (1927). *Studies in optics*. University of Chicago Press.

Michna, M. L., Yoshizawa, T., & Mullen, K. T. (2007). S-cone contributions to linear and non-linear motion processing. *Vision Research*, *47*(8), 1042-1054. (6th Annual Meeting of the Vision-Sciences-Society, Sarasota, FL, 2006)

Mollon, J. D. (2003). Color science. In S. Shevell (Ed.), (p. 1-39). Optical Society of America.

Morgan, M. J., Adam, A., & Mollon, J. D. (1992). Dichromats detect colour-camouflaged objects that are not detected by trichromats. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *248*, 291–295.

Morrone, M. C., Burr, D. C., & Maffei, L. (1982). Functional implications of cross-orientation inhibition of cortical visual cells. i. neurophysiological evidence. *Proceedings of the Royal Society of London B*, *216*(1204), 335–354.

Mullen, K. T. (1985). The contrast sensitivity of human colour-vision to red green and blue yellow chromatic gratings. *The Journal of Physiology*, *359*, 381–400.

Mullen, K. T., & Beaudot, W. H. A. (2002). Comparison of color and luminance vision on a global shape discrimination task. *Vision Research*, *42*, 565–575.

Mullen, K. T., Beaudot, W. H. A., & McIlhagga, W. H. (2000). Contour integration in color vision: a common process for the blueyellow, redgreen and luminance mechanisms? *Vision Research*, *40*, 639–655.

Mullen, K. T., Gheiratmand, M., Medina, J. M., & Kim, Y. J. (2012). Two routes to suppression of signals in color vision. *Journal of Vision*, *12*(9).

Mullen, K. T., & Losada, M. A. (1994). Evidence for separate pathways for color and luminance detection mechanisms. *Journal of the Optical Society of America A - Optics, Image Science and Vision*, *11*(12), 3136–3151.

Mullen, K. T., Thompson, B., & Hess, R. F. (2010). Responses of the human visual cortex and LGN to achromatic and chromatic temporal modulation: an fMRI study. *Journal of Vision*, *10*(13).

Nachmias, J., & Sansbury, R. V. (1974). Grating contrast: discrimination may be better than detection. *Vision Research*, *14*(10), 1039–1042.

Naka, K. I., & Rushton, W. A. H. (1966). S-potentials from luminosity units in the retina of fish (cyprinidae). *Journal of Physiology*, *1855*, 587–599.

Nayatani, Y. (1998). A colorimetric explanation of the Helmholtz-Kohlrausch effect. *COLOR Research and Application*, *23*, 374–378.

Nefs, H. T., Koenderink, J. J., & Kappers, A. M. L. (2005). The influence of illumination direction on the pictorial reliefs of lambertian surfaces. *Perception*, *34*(3), 275–287.

Nefs, H. T., O'Hare, L., & Harris, J. M. (2010). Two independent mechanisms for motion-in-depth perception: Evidence from individual differences. *Frontiers in Psychology*, *1*(13).

Neitz, J., & Neitz, M. (2011). The genetics of normal and defective color vision. *Vision Research*, *51*, 631–651.

Nelson, R., & Kolb, H. (2004). ON and OFF pathways in the vertebrate retina and

visual system. In L. M. Chalupa & J. S. Werner (Eds.), *The visual neurosciences* (p. 260-278). The MIT Press.

Nguyen-Tri, D., Overbury, O., & Faubert, J. (2003). The role of lenticular senescence in age-related color vision changes. *Visual Psychophysics and Physiological Optics*, *44*(8), 3698–3704.

Norcia, A. (2004). The visual neurosciences. In L. M. Chalupa & J. S. Werner (Eds.), (pp. 174–188). Cambridge, MA: The MIT Press.

Olmos, A., & Kingdom, F. A. A. (2004). A biologically inspired algorithm for the recovery of shading and reflectance images. *Perception*, *33*(12), 1463–1473.

Parry, N. R. A., Murray, I. J., Panorgias, A., McKeefry, D. J., Lee, B. B., & Kremers, J. (2012). Simultaneous chromatic and luminance human electroretinogram responses. *The Journal of Physiology*, *590*, 3141–3154.

Parry, N. R. A., & Robson, A. G. (2012). Optimization of large field tritan stimuli using concentric isoluminant annuli. *Journal of Vision*, *12*(12).

Pearson, P., & Kingdom, F. A. A. (2002). Texture-orientation mechanisms pool colour and luminance contrast. *Vision Research*, *42*(12), 1547-1558.

Peli, E. (1990). Contrast in complex images. *Journal of the Optical Society of America A - Optics, Image Science and Vision*, *7*(10), 2032–2040.

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

Pentland, A. P. (1982a). Finding the illuminant direction. *Journal of the Optical Society of America*, *72*(4), 448–455.

Pentland, A. P. (1982b). The perception of shape from shading. In *Proceedings of the osa annual meeting.*

Petrov, Y., Carandini, M., & McKee, S. (2005). Two distinct mechanisms of suppression in human vision. *The Journal of Neuroscience*, *25*(38), 8704–8707.

Phillips, G. C., & Wilson, H. R. (1984). Orientation bandwidths of spatial mechanisms measured by masking. *Journal of the Optical Society of America*, *1*(2), 226–232.

Planck, M. (1914). *The theory of heat radiation* (M. Masius, Trans.).

Poggio, G., Baker, F., Mansfield, R., Sillito, A., & Grigg, P. (1975). Spatial and chromatic properties of neurons subserving foveal and parafoveal vision in rhesus monkey. *Brain Research*, *100*(1), 25–59.

Pokorny, J., Smith, V. C., & Lutze, M. (1989). Hetereochromatic modulation photometry. *Journal of the Optical Society of America*, *6*(10), 1618–1623.

Polat, U., & Norcia, A. M. (1998). Elongated physiological summation pools in the human visual cortex. *Vision Research*, *38*(23), 3735–3741.

Polyak, S. L. (1941). *The retina.* University of Chicago Press.

Polyak, S. L. (1957). *The vertebrate visual system.* University of Chicago Press.

Potetz, B., & Lee, T. S. (2003). Statistical correlations between two-dimensional images and three-dimensional structures in natural scenes. *Journal of the Optical Society of America*, *20*(7), 1292–1303.

Priebe, N., & Ferster, D. (2006). Mechanisms underlying cross-orientation suppression in cat visual cortex. *Nature Neuroscience*, *9*(4), 552–561.

Quick, R. F., Hamerly, J. R., & Reichert, T. A. (1976). Absence of a measurable critical band at low suprathreshold contrasts. *Vision Research*, *16*(4), 351–355.

Ramachandran, V. S. (1988). Perception of shape from shading. *Nature*, *331*(6152), 163–166.

Ramachandran, V. S. (1991). Interactions between motion, depth, color and form: the utilitarian theory of perception. In C. Blakemore (Ed.), *Vision, coding and efficiency* (pp. 346–360). Cambridge University Press.

Ramon y Cajal, S., & Felipe, S. (1892). La rétine des vertébrés. *La cellule*, *1*(9).

Regan, D. (1991). Spatial vision for objects defined by colour contrast, binocular disparity and motion parallax. In D. Regan (Ed.), *Spatial vision* (pp. 135–178). MacMillian Press.

Regan, D., & Regan, M. P. (1987). Nonlinearity in human visual responses to two-dimensional patterns, and a limitation of fourier methods. *Vision Research*, *27*(12), 2181–2183.

Regan, D., & Tyler, C. (1971). Some dynamic features of colour vision. *Vision Research*, *11*(11), 1307–1324.

Reisbeck, T. E., & Gegenfurtner, K. R. (1998). Effects of contrast and temporal frequency on orientation discrimination for luminance and isoluminant stimuli. *Vision Research*, *38*(8), 1105–1117.

Ripamonti, C., Woo, W. L., Crowther, E., & Stockman, A. (2009). The S-cone contribution to luminance depends on the M- and L-cone adaptation levels: Silent surrounds? *Journal of Vision*, *9*(3).

Robinson, A., & MacLeod, D. (2011). The McCollough effect with plaids and gratings: Evidence for a plaid-selective visual mechanism. *Journal of Vision*, *11*(26).

Rodieck, R. W. (1965). Quantitative analysis of cat retinal ganglion cell response to visual stimuli. *Vision Research*, *5*(11), 583–601.

Rodieck, R. W., & Stone, J. (1965). Analysis of receptive fields of cat retinal ganglion cells. *Journal of Neurophysiology*, *28*, 833–849.

Roorda, A., Metha, A. B., Lennie, P., & Williams, D. R. (2001). Packing arrangement of the three cone classes in primate retina. *Vision Research*, *41*, 1291–1306.

Ross, J., & Speed, H. D. (1991). Contrast adaptation and contrast masking in human vision. *Proceedings of the Royal Society of London B*, *246*(1315), 61–69.

Rubin, J. M., & Richards, W. A. (1982). Color-vision and image intensities: When are changes material. *Biological Cybernetics*, *45*(3), 215–226.

Sanchez, J. M., & Fairchild, M. D. (2002). Quantification of the Helmholtz-Kohlrausch effect for CRT color monitors. In *Spie proceedings 4421, 9th congress of the international colour association.*

Sanders, C. L., & Wyszecki, G. (1957). Correlate for lightness in terms of CIE-tristimulus values. Part I. *Journal of the Optical Society of America*, *47*(5), 398–404.

Sanders, C. L., & Wyszecki, G. (1958). L/Y ratios in terms of CIE-chromaticity coordinates. *Journal of the Optical Society of America*, *48*(6), 389–391.

Schade, O. H. (1956). Optical and photoelectric analog of the eye. *Journal of the Optical Society of America*, *46*(9), 721–739.

Schnapf, J. L., Nunn, B. J., Meister, M., & Baylor, D. A. (1990). Visual transduction in cones of the monkey macaca fascicularis. *Journal of Physiology*, *427*, 681–713.

Schofield, A. J., & Georgeson, M. A. (1999). Sensitivity to modulations of luminance and contrast in visual white noise: separate mechanisms with similar behaviour. *Vision Research*, *39*(16), 2697–2716.

Schofield, A. J., Hesse, G., Rock, P. B., & Georgeson, M. A. (2006). Local luminance amplitude modulates the interpretation of shape-from-shading in textured surfaces. *Vision Research*, *46*(20), 3462–3482.

Schofield, A. J., Rock, P. B., Sun, P., Jiang, X., & Georgeson, M. A. (2010). What is second-order vision for? Discriminating illumination versus material changes. *Journal of Vision*, *10*(9).

Schofield, A. J., Sun, P., & Mazzilli, G. (2013). Shape perception in human and computer vision. In S. Dickinson & Z. Pizlo (Eds.), *Advances in computer vision and pattern recognition* (pp. 119–132). Springer-Verlag London.

Sclar, G., Maunsell, J. H. R., & Lennie, P. (1990). Coding of image contrast in the central visual pathway of the macaque monkey. *Vision Research*, *30*(1), 1–10.

Seriès, P., Lorenceau, J., & Frégnac, Y. (2003). The "silent" surround of V1 receptive fields: theory and experiments. *Journal of Physiology*, *97*(4–6), 453–47.

Shapley, R., & Enroth-Cugell, C. (1984). Progress in retinal research. In O. N. & C. G. (Eds.), (pp. 263–346). Pergamon, London.

Shapley, R., & Hawken, M. J. (2011). Color in the cortex: single- and double-opponent cells. *Vision Research*, *51*, 701–717.

Sharifian, F., Nurminen, L., & Vanni, S. (2013). Visual interactions conform to pattern decorrelation in multiple cortical areas. *PLoS ONE*, *8*(7).

Sharpe, L. T., de Luca, E., Hansen, T., Jägle, H., & Gegenfurtner, K. R. (2006). Advantages and disadvantages of human dichromacy. *Journal of Vision*, *6(3)*.

Sharpe, L. T., Stockman, A., Jägle, H., Knau, H., Klausen, G., Reitner, A., & Nathans, J. (1998). Red, green, and red-green hybrid pigments in the human retina: Correlations between deduced protein sequences and psychophysically measured spectral sensitivities. *The Journal of Neuroscience*, *18(23)*, 10053–10069.

Sherman, S. M., & Guillery, R. W. (1996). Functional organization of thalamocortical relays. *Journal of Neurophysiology*, *76*(3), 1367–1395.

Shevell, S. K., & Kingdom, F. A. A. (2008). Color in complex scenes. *Annual Review of Psychology*, *59*, 143–166.

Shiells, R. A., Falk, G., & Naghshineh, S. (1981). Action of glutamate and aspartate analogues on rod horizontal and bipolar cells. *Nature*, *294*(5841), 592–594.

Simmons, D. R., & Kingdom, F. A. A. (1997). On the independence of chromatic and

achromatic stereopsis mechanisms. *Vision Research*, *37*(10), 1271–1280.

Simmons, D. R., & Kingdom, F. A. A. (2002). Interactions between chromatic- and luminance-contrast-sensitive stereopsis mechanisms. *Vision Research*, *42*(12), 1535–1545.

Singer, B., & D'Zmura, M. (1994). Color contrast induction. *Vision Research*, *34*(23), 3111–3126.

Smith, D. P., Cole, B. L., & Isaac, A. (1973). Congenital tritanopia without neuroretinal disease. *Investigative Ophtamology*, *12(8)*, 608–617.

Smith, M. A., Kelly, R. C., & Lee, T. S. (2007). Dynamics of response to perceptual pop-out stimuli in macaque V1. *Journal of Neurophysiology*, *98*, 3436–3449.

Smith, V. C., & Pokorny, J. (1975). Spectral sensitivity of the foveal cone photopigments between 400 and 500 nm. *Vision Research*, *15*, 161–171.

Solomon, J. A. (2009). The history of dipper functions. *Attention, Perception, & Psychophysics*, *71*(3), 435–443.

Solomon, J. A., Sperling, G., & Chubb, C. (1993). The lateral inhibition of perceived contrast is indifferent to on-center/off-center segregation, but specific to orientation. *Vision Research*, *33*(18), 2671–2683.

Solomon, S. G. (2002). Striate cortex in dichromatic and trichromatic marmosets: Neurochemical compartmentalization and geniculate input. *The Journal of Comparative Neurology*, *450*, 366–381.

Solomon, S. G., & Lennie, P. (2005). Chromatic gain controls in visual cortical neurons. *The Journal of Neuroscience*, *25*(19), 4779–4792.

Solomon, S. G., & Lennie, P. (2007). The machinery of colour vision. *Nature Reviews Neuroscience*, *8*(4), 276–286.

Speed, H. D., & Ross, J. (1992). Spatial frequency tuning of facilitation by masks. *Vision Research*, *32*(6), 1143–1148.

Sterling, P. (1999). The ganglion cell receptive field. In J. I. Toyoda, M. Murakami, A. Kaneko, & T. Saito (Eds.), *The retinal basis of vision* (pp. 163–169). Elsevier Science.

Stockman, A., & Brainard, D. H. (2010). Color vision mechanisms. In M. Bass (Ed.), *Osa handbook of optics* (3rd ed., pp. 11.1–11.104). McGraw-Hill, New York.

Stockman, A., Langendörfer, M., Smithson, H. E., & Sharpe, L. T. (2006). Human cone light adaptation: From behavioral measurements to molecular mechanisms. *Journal of Vision*, *6*(11).

Stockman, A., & Sharpe, L. (2000). The spectral sensitivities of the middle-and long-wavelength-sensitive cones derived from measurements in observers of known genotype. *Vision Research*, *40*(13), 1711–1737.

Stromeyer, C. F., & Klein, S. (1974). Spatial frequency channels in human vision as asymmetric (edge) mechanisms. *Vision Research*, *14*, 1409–1420.

Stromeyer, C. F., Thabet, R., Chaparro, A., & Kronauer, R. E. (1999). Spatial masking does not reveal mechanisms selective to combined luminance and red-green color.

Vision Research, *39*(12), 2099–2112.

Sun, H., Smithson, H. E., Zaidi, Q., & Lee, B. B. (2006). Do magnocellular and parvocellular ganglion cells avoid short-wavelength cone input? *Visual Neuroscience*, *23*(3-4), 441–446.

Sun, J., & Perona, P. (1998). Where is the sun? *Nature Neuroscience*, *1*(3), 183–184.

Sun, P., & Schofield, A. J. (2012). Two operational modes in the perception of shape from shading revealed by the effects of edge information in slant settings. *Journal of Vision*, *12*(1).

Swanson, W. H., Ueno, T., Smith, V. C., & Pokorny, J. (1987). Temporal modulation sensitivity and pulse-detection thresholds for chromatic and luminance perturbations. *Journal of the Optical Society of America*, *4*(10), 1992–2005.

Swift, D. J., & Smith, R. A. (1983). Spatial frequency masking and Webers law. *Vision Research*, *23*(5), 495–505.

Switkes, E., Bradley, A., & Devalois, K. (1988). Contrast dependence and mechanisms of masking interactions among chromatic and luminance gratings. *Journal of the Optical Society of America A - Optics, Image Science and Vision*, *5*(7), 1149-1162.

Switkes, E., & Crognale, M. A. (1999). Comparison of color and luminance contrast: Apples versus oranges? *Vision Research*, *39*(10), 1823–1831.

Szmajda, B. A., Buzás, P., FitzGibbon, T., & Martin, P. R. (2006). Geniculocortical relay of blue-off signals in the primate visual system. *Proceedings of the National Academy of Sciences*, *103*(51), 19512–19517.

Szmajda, B. A., Grünert, U., & Martin, P. R. (2008). Retinal ganglion cell inputs to the koniocellular pathway. *The Journal of Comparative Neurology*, *510*, 251–268.

Tailby, C., Solomon, S. G., Dhruv, N. T., & Lennie, P. (2008). Habituation reveals fundamental chromatic mechanisms in striate cortex of macaque. *Journal of Neuroscience*, *28*(5), 1131–1139.

Tailby, C., Solomon, S. G., & Lennie, P. (2008). Functional asymmetries in visual pathways carrying S-cone signals in macaque. *Journal of Neuroscience*, *28*(15), 4078–4087.

Tajima, S., Watanabe, M., Imai, C., Ueno, K., Asamizuya, T., Sun, P., ... Cheng, K. (2010). Opposing effects of contextual surround in human early visual cortex revealed by functional magnetic resonance imaging with continuously modulated visual stimuli. *The Journal of Physiology*, *30*(9), 3264–3270.

Thorell, L. G., de Valois, R. L., & Albrecht, D. G. (1984). Spatial mapping of monkey V1 cells with pure color and luminance stimuli. *Vision Research*, *24*(7), 751–769.

Tiippana, K., & Näsänen, R. (1999). Spatial-frequency bandwidth of perceived contrast. *Vision Research*, *39*(20), 3399–3403.

Tiippana, K., Rovamo, J., Näsänen, R., Whitaker, D., & Mäkelä, P. (2000). Contrast matching across spatial frequencies for isoluminant chromatic gratings. *Vision Research*, *40*(16), 2159–2165.

Todd, J. T., & Mingolla, E. (1983). Perception of surface curvature and direction of illumination from patterns of shading. *Journal of Experimental Psychology: Human*

Perception and Performance, *9*(4), 583–595.

Tyler, C. W. (1998). Diffuse illumination as a default assumption for shape-from-shading in the absence of shadows. *Journal of Imaging Science and Technology*, *42*(4), 319–325.

Victor, J. D., Purpura, K. P., & Conte, M. M. (1998). Chromatic and luminance interactions in spatial contrast signals. *Visual Neuroscience*, *15*(4), 607–624. (67th Annual Meeting of the Association-for-Research-in-Vision-and-Ophthalmology, FT LAUDERDALE, FLORIDA, APR 21-26, 1996)

Vimal, R. L. P. (1997). Orientation tuning of the spatial-frequency-tuned mechanisms of the redgreen channel. *Journal of the Optical Society of America*, *14*(10), 2622–2632.

Vimal, R. L. P. (1998a). Color-luminance interaction: Data produced by oblique cross masking. *Journal of the Optical Society of America*, *15*(7), 1756–1766.

Vimal, R. L. P. (1998b). Spatial-frequency tuning of sustained nonoriented units of the redgreen channel. *Journal of the Optical Society of America*, *15*(1), 1–15.

Vimal, R. L. P. (2000). Spatial color contrast matching: broad-bandpass functions and the flattening effect. *Vision Research*, *40*(23), 3231–3243.

Vimal, R. L. P. (2002). Spatial-frequency-tuned mechanisms of the red-green channel estimated by oblique masking. *Journal of the Optical Society of America*, *19*(2), 276–288.

von Békésy. (1960). *Experiments in hearing* (E. G. Wever, Trans.). New York: McGraw-Hill.

von Helmholtz, H. (1852). Über die Theorie der zusammengesetzten Farben. *Annalen der Physik*, *87*, 45–66.

von Helmholtz, H. (1867). *Handbuch der physiologischen Optik*. Leipzig: Leopold Voss.

Vos, J. (1978). Colorimetric and photometric properties of a 2° fundamental observer. *Color Research & Application*, *3*, 125–128.

Wachtler, T., Sejnowski, T. J., & Albright, T. D. (2003). Representation of color stimuli in awake macaque primary visual cortex. *Neuron*, *37*, 681–691.

Wade, A., Augath, M., Logothetis, N., & Wandell, B. (2008). fMRI measurements of color in macaque and human. *Journal of Vision*, *8*(10).

Wagner, G., & Boynton, R. M. (1972). Comparison of four methods of heterochromatic photometry. *Journal of the Optical Society of America*, *62*(12), 1508–1515.

Watson, A. B., & Solomon, J. A. (1997). Model of visual contrast gain control and pattern masking. *Journal of the Optical Society of America*, *14*(9), 2379–2391.

Webster, M. A. (1996). Human colour perception and its adaptation. *Network: Computation in Neural Systems*, *7*, 587–634.

Webster, M. A., DeValois, K. K., & Switkes, E. (1990). Orientation and spatial-frequency discrimination for luminance and chromatic gratings. *Journal of the Optical Society of America*, *7*, 1034–1049.

Webster, M. A., & Malkoc, G. (2000). Color-luminance relationships and the McCollough effect. *Perception and Psychophysics*, *62*(4), 659–672.

Webster, M. A., & Mollon, J. D. (1993). Contrast adaptation dissociates different measures of luminous efficiency. *Journal of the Optical Society of America*, *10*(6), 1332–1340.

Werner, A. (2003). The spatial tuning of chromatic adaptation. *Vision Research*, *43*, 1611–1623.

Werner, A., Sharpe, L. T., & Zrenner, E. (2000). Asymmetries in the time-course of chromatic adaptation and the significance of contrast. *Vision Research*, *40*(9), 1101–1113.

Wiesel, T. N., & Hubel, D. H. (1966). Spatial and chromatic interactions in the lateral geniculate body of the rhesus monkey. *Journal of Neurophysiology*, *29*(6), 1115–1156.

Williams, A. L., Singh, K. D., & Smith, A. T. (2003). Surround modulation measured with functional mri in human visual cortex. *Journal of Neurophysiology*, *89*, 525–533.

Wilson, H. R. (1980). A transducer function for threshold and suprathreshold human vision. *Biological Cybernetics*, *38*, 171–178.

Wilson, H. R. (1990). Psychological models of spatial visionand hyperacuity. In D. Regan (Ed.), *Spatial vision* (pp. 64–86). MacMillian Press.

Wilson, H. R., & Wilkinson, F. (2004). Spatial channels in vision and spatial pooling. In L. M. Chalupa & J. S. Werner (Eds.), *The visual neuroscience* (pp. 1060–1068). Cambridge: The MIT Press.

Withouck, M., Smet, K. A. G., Ryckaert, W. R., Pointer, M. R., Deconinck, G., Koenderink, J., & Hanselaer, P. (2013). Brightness perception of unrelated self-luminous colors. *Journal of the Optical Society of America*, *30*(6), 1248–1255.

Wright, M. D. (1946). *Researches on normal and defective colour vision.* London: Kimpton.

Wright, M. J., & Ledgeway, T. (2004). Interaction between luminance gratings and disparity gratings. *Spatial Vision*, *17*(1-2), 51–74.

Wright, W. D. (1928). A trichromatic colorimeter with spectral primaries. *Transactions of the Optical Society*, *29*, 225-243.

Wuerger, S. M., & Landy, M. S. (1993). Role of chromatic and luminance contrast in inferring structure from motion. *Journal of the Optical Society of America*, *10*(6), 1363–1372.

Wyszecki, G. (1965). Matching color differecnes. *Journal of the Optical Society of America*, *55*(10), 1319–1324.

Wyszecki, G. (1967). Correlate for lightness in terms of CIE chromaticity coordinates and luminous reflectance. *Journal of the Optical Society of America*, *57*(2), 254–257.

Wyszecki, G., & Sanders, C. L. (1957). Correlate for lightness in terms of CIE-tristimulus values. part II. *Journal of the Optical Society of America*, *47*(9), 840–842.

Wyszecki, G., & Stiles, W. S. (1982). *Color science: Concepts and methods, quantitative data and formulae* (Second ed.). New York: Wiley.

Xing, D., Shapley, R. M., Hawken, M. J., & Ringach, D. L. (2005). Effect of stimulus size on the dynamics of orientation selectivity in macaque V1. *Journal of Neurophysiology*,

94, 799–812.

Xing, J., & Heeger, D. J. (2000). Center-surround interactions in foveal and peripheral vision. *Vision Research*, *40*(22), 3065–3072.

Yau, K.-W. (1994). Phototransduction mechanism in retinal rods and cones. The Friedenwald lecture. *Investigative Ophthalmology & Visual Science*, *35*(1), 9–32.

Young, T. (1802). The Bakerian lecture. On the theory of light and colours. *Philosophical Transactions of the Royal Society of London*, *92*, 12–48.

Yu, C., Klein, S. A., & Levi, D. M. (2001). Surround modulation of perceived contrast and the role of brightness induction. *Journal of Vision*, *1*(1).

Yu, C., Klein, S. A., & Levi, D. M. (2002). Facilitation of contrast detection by cross-oriented surround stimuli and its psychophysical mechanisms. *Journal of Vision*, *2*, 243–255.

Yu, C., & Levi, D. M. (2000). Surround modulation in human vision unmasked by masking experiments. *Nature Neuroscience*, *3*(7), 724–728.

Zaidi, Q., Spehar, B., & DeBonet, J. (1997). Color constancy in variegated scenes: role of low-level mechanisms in discounting illumination changes. *Journal of the Optical Society of America*, *14*(10), 2608–2621.

Appendix A

Ethics forms

University of St Andrews

University Teaching and Research Ethics Committee
School of Psychology

15 February 2010

Ethics Reference No: *Please quote this ref on all correspondence*	PS6135
Project Title:	Visual Cues to Shape and Depth Perception
Researchers Name(s):	Stephane Clery, Dr George Lovell
Supervisor(s):	Professor Julie Harris

Thank you for submitting your application which was considered at the School Ethics Committee meeting on the 12th February 2010. The following documents were reviewed:-

1. Ethical Application Form 12/02/2010
2. Participant Information Sheet 12/02/2010
3. Consent Form 12/02/2010
4. Debriefing Form 12/02/2010
5. Advertisement 12/02/2010

The University Teaching and Research Ethics Committee (UTREC) approves this study from an ethical point of view. Please note that where approval is given by a School Ethics Committee that committee is part of UTREC and is delegated to act for UTREC.

Approval is given for three years. Projects, which have not commenced within two years of original approval, must be re-submitted to your School Ethics Committee.

You must inform your School Ethics Committee when the research has been completed. If you are unable to complete your research within the 3 three year validation period, you will be required to write to your School Ethics Committee and to UTREC (where approval was given by UTREC) to request an extension or you will need to re-apply.

Any serious adverse events or significant change which occurs in connection with this study and/or which may alter its ethical consideration, must be reported immediately to the School Ethics Committee, and an *Ethical* Amendment Form submitted where appropriate.

Approval is given on the understanding that the 'Guidelines for Ethical Research Practice' (http://www.st-andrews.ac.uk/media/UTRECguidelines%20Feb%2008.pdf) are adhered to.

Yours sincerely

Convenor of the School Ethics Committee OR Convener of UTREC

Ccs Professor Julie Harris (Supervisor)
School Ethics Committee

8 April 2011

Ethics Reference No: *Please quote this ref on all correspondence*	PS6135 (Amendment)
Project Title:	Visual cues to shape and depth perception
Researchers Name(s):	Stephane Clery, Dr George Lovell, Anastasia Kolokolnikova
Supervisor(s):	Professor Julie Harris

Thank you for submitting your application which was considered at the Psychology School Ethics Committee meeting on the 8th April 2011. The following documents were reviewed:

1. Ethical Amendment Form 08/04/2011
2. Participant Information Sheet 08/04/2011
3. Consent Form 08/04/2011
4. Debriefing Form 08/04/2011

The University Teaching and Research Ethics Committee (UTREC) approves this study from an ethical point of view. Please note that where approval is given by a School Ethics Committee that committee is part of UTREC and is delegated to act for UTREC.

Approval is given for three years. Projects, which have not commenced within two years of original approval, must be re-submitted to your School Ethics Committee.

You must inform your School Ethics Committee when the research has been completed. If you are unable to complete your research within the 3 three year validation period, you will be required to write to your School Ethics Committee and to UTREC (where approval was given by UTREC) to request an extension or you will need to re-apply.

Any serious adverse events or significant change which occurs in connection with this study and/or which may alter its ethical consideration, must be reported immediately to the School Ethics Committee, and an Ethical Amendment Form submitted where appropriate.

Approval is given on the understanding that the 'Guidelines for Ethical Research Practice' (http://www.st-andrews.ac.uk/media/UTRECguidelines%20Feb%2008.pdf) are adhered to.

Yours sincerely

Convenor of the School Ethics Committee

Ccs Prof. Julie Harris (Supervisor)
School Ethics Committee

Appendix B

Published paper

This paper has been removed due to copyright restrictions. Clery et al. (2013). doi: 10.1167/13.5.16