INTEROCULAR SUPPRESSION AND CONTRAST GAIN CONTROL IN HUMAN VISION

Daniel Hart Baker

Doctor of Philosophy

ASTON UNIVERSITY

February 2008

Aston University

Interocular suppression and contrast gain control in human vision

Daniel Hart Baker

Doctor of Philosophy

2007

Summary

The human visual system combines contrast information from the two eyes to produce a single cyclopean representation of the external world. This task requires both summation of congruent images and inhibition of incongruent images across the eyes. These processes were explored psychophysically using narrowband sinusoidal grating stimuli. Initial experiments focussed on binocular interactions within a single detecting mechanism, using contrast discrimination and contrast matching tasks. Consistent with previous findings, dichoptic presentation produced greater masking than monocular or binocular presentation. Four computational models were compared, two of which performed well on all data sets. Suppression between mechanisms was then investigated, using orthogonal and oblique stimuli. Two distinct suppressive pathways were identified, corresponding to monocular and dichoptic presentation. Both pathways impact prior to binocular summation of signals, and differ in their strengths, tuning, and response to adaptation, consistent with recent single-cell findings in cat. Strikingly, the magnitude of dichoptic masking was found to be spatiotemporally scale invariant, whereas monocular masking was dependent on stimulus speed. Interocular suppression was further explored using a novel manipulation, whereby stimuli were presented in dichoptic antiphase. Consistent with the predictions of a computational model, this produced weaker masking than in-phase presentation. This allowed the bandwidths of suppression to be measured without the complicating factor of additive combination of mask and test. Finally, contrast vision in strabismic amblyopia was investigated. Although amblyopes are generally believed to have impaired binocular vision, binocular summation was shown to be intact when stimuli were normalized for interocular sensitivity differences. An alternative account of amblyopia was developed, in which signals in the affected eye are subject to attenuation and additive noise prior to binocular combination.

Key words:

Visual psychophysics;
Suppression;
Contrast gain control;
Dichoptic masking;
Amblyopia

For Laura

Acknowledgements

As my supervisor, Tim Meese has been an invaluable source of knowledge, guidance, clarity and wisdom. I can imagine no better introduction to the world of academia and vision research than that which I have gained through him. His enthusiasm for the subject is unparalleled, despite his successful avoidance of sitting as a subject in any of my experiments! There is no way that my PhD experience could have been anywhere near as enjoyable without his involvement, so thank you Tim for this marvellous opportunity.

Mark Georgeson, as my associate supervisor, has been perpetually available for advice and support. Particular thanks to Mark for getting me interested in signal detection theory, for extensive discussions on binocular rivalry, and for finally admitting that it is possible for me to exist without caffeine!

I am grateful to Robert Hess for lab space, facilities, ideas and encouragement during my various trips to Montréal. Robert was an ideal external supervisor, combining prolific inventiveness with an encyclopaedic knowledge of the field. Further thanks to all at the McGill Vision Research unit for making me feel welcome during my visits.

Rob Summers, by allowing me to use his fantastic 'Liberator' software, has saved me literally hundreds of hours which I would otherwise have spent writing code to run experiments. Rob was always on hand to incorporate a new feature, or fix a bug, and this has played a major role in the abundance of experiments I was able to run. He was also responsible for many enjoyable pub visits during his final months of freedom before becoming a father.

I am greatly indebted to all those who helped out by serving as experimental observers, and wish to apologise profusely to all of you! Special commendations for psychophysical hardiness go to Laura Patryas, Stuart Wallis, Tim Yates, and Rob Summers, all of whom gave up many hours of their time to gather data. Thanks guys, it was all worth it really!

A huge thank you also goes to all my friends, especially my housemates, Guy, Jack, Rox and Helena for general support and many poker evenings, Chani Bond for coming out to Canada, and Ally Donaldson for lengthy and impenetrable phone calls about something called 'model checking', which I still can barely comprehend. Thanks also to my parents for general support. Finally, a general thank you to all in the Aston vision lab and its environs, particularly Kirsten Challinor for her fabulous cooking and baking skills.

The work contained in this thesis is my own, and is presented in this form for the first time. There is one exception to this, regarding two computational models presented in Chapter 3 which were formulated by my supervisor, Dr Tim Meese. However, the implementation of these models is my own, as is all of the accompanying text. This contribution is also clearly indicated in the text.

Signed

Daniel Baker

"You'll keep an eye on Frodo, won't you?"

"Yes, I will - two eyes, as often as I can spare them."

- The Fellowship of the Ring, J.R.R. Tolkien

Contents

1 Introduction **18**

1.1 Binocular vision 18

1.2 Important findings in contrast perception 19

1.2.1 Detection thresholds and contrast sensitivity 19

1.2.2 Within-channel masking 19

1.2.3 Modelling the dipper function 20

1.2.4 Multiple channels in human vision 22

1.2.5 Cross-channel masking 23

1.3 Contrast gain control 24

1.4 Presenting stimuli to different eyes 25

1.4.1 Binocular summation at threshold 26

1.4.2 Binocular summation above threshold 28

1.4.3 Interocular contrast matching 28

1.4.4 Dichoptic masking and binocular rivalry 30

1.5 Models of binocular interaction 31

1.6 Contrast vision in amblyopia 32

1.7 Overview 33

1.8 Collaborative work and publication 35

2 General methods **37**

2.1 Using psychophysics to explore visual processing 37

2.2 Stimuli . . . 38
2.3 Frame interleaving and contrast scaling . . . 39
2.4 Equipment . . . 40
2.4.1 Computing hardware . . . 40
2.4.2 Monitors . . . 40
2.4.3 Shutter goggles . . . 41
2.4.4 Mirror stereoscope . . . 41
2.5 Experimental software . . . 42
2.6 Data analysis and modelling . . . 42
2.7 Principal Observer . . . 43
2.8 Procedure . . . 43

3 Within-channel masking **46**

3.1 Experiment 1: dipper functions . . . 46
3.1.1 Introduction . . . 46
3.1.2 Methods . . . 46
3.1.3 Results . . . 47
3.1.4 Models . . . 50
3.1.5 Predicting psychometric slopes . . . 56
3.1.6 Surviving models . . . 57
3.2 Experiment 2: twin mask contrast discrimination . . . 59
3.2.1 Introduction . . . 59
3.2.2 Methods . . . 60
3.2.3 Results . . . 60
3.2.4 Models . . . 62
3.3 Experiment 3: contrast matching . . . 63
3.3.1 Introduction . . . 63
3.3.2 Methods . . . 64

3.3.3 Results . . . 65
3.3.4 Models of matching . . . 67
3.3.5 Discussion . . . 67
3.4 Experiment 4: paradoxical psychometric functions . . . 68
3.4.1 Introduction . . . 68
3.4.2 Methods . . . 69
3.4.3 Results . . . 69
3.4.4 Discussion . . . 70
3.5 Comparison of two models . . . 71
3.6 Chapter summary . . . 71

4 Initial investigations into interocular cross-channel masking 73

4.1 Chapter Introduction . . . 73
4.2 Experiment 5: Cross-channel masking in five ocular configurations . . . 74
4.2.1 Introduction . . . 74
4.2.2 Methods . . . 75
4.2.3 Results . . . 76
4.3 Experiment 6: Cross-channel masking at varying stimulus durations . . . 80
4.3.1 Methods . . . 80
4.3.2 Results . . . 80
4.4 Summary . . . 82
4.5 Extending the Two Stage Model . . . 82
4.5.1 Basic model . . . 82
4.5.2 Describing subtleties of the data set: Models 2 & 3 . . . 86
4.5.3 Temporal components . . . 89
4.6 Extending the Twin Summation Model . . . 92
4.7 Conclusions . . . 93

5 Further cross-channel masking experiments, influenced by neurophysiol-

ogy **94**

5.1 Physiology of cross-channel suppression 94

5.2 Experiment 7 - cross-channel twin mask experiment 96

5.2.1 Methods 96

5.2.2 Results 97

5.2.3 Models 98

5.2.4 Discussion 102

5.3 Experiment 8 - adaptation of cross-channel masking 103

5.3.1 Methods 104

5.3.2 Results 104

5.3.3 Discussion 105

5.3.4 Modelling 107

5.4 Biologically plausible models 110

5.4.1 Monocular masking 111

5.4.2 Feedback models 112

5.5 Experiment 9 - spatiotemporal tuning of orthogonal masking 116

5.5.1 Methods 116

5.5.2 Results 118

5.5.3 Discussion 120

5.6 Experiment 10 - scale dependence of cross-channel masking 121

5.6.1 Methods 121

5.6.2 Results 122

5.6.3 Modelling 126

5.6.4 Discussion 128

5.7 Chapter Summary 134

6 Experiments using stimuli in dichoptic antiphase **135**

6.1 Experiment 11: dichoptic masking with antiphase stimuli 135

6.1.1 Introduction . . . 135
6.1.2 Methods . . . 136
6.1.3 Results . . . 137
6.1.4 Modelling - the nature of dichoptic masking . . . 137
6.1.5 Phase tuning of dichoptic masking . . . 139
6.2 Experiment 12: dichoptic spatial frequency tuning . . . 141
6.2.1 Introduction - Legge's unusual masking functions . . . 141
6.2.2 Methods . . . 142
6.2.3 Results . . . 143
6.2.4 Discussion . . . 144
6.3 Experiment 13: dichoptic orientation tuning . . . 145
6.3.1 Introduction . . . 145
6.3.2 Methods . . . 146
6.3.3 Results . . . 146
6.3.4 Discussion . . . 147
6.4 Modelling dichoptic masking functions . . . 147
6.4.1 Orientation model . . . 148
6.4.2 Spatial Frequency model . . . 153
6.5 Monocular tuning functions . . . 155
6.5.1 Results . . . 156
6.5.2 Comparison with dichoptic model . . . 158
6.6 Chapter discussion . . . 159
6.6.1 Predictions of the dichoptic model . . . 160

7 Amblyopia **164**

7.1 Amblyopic contrast vision . . . 164
7.2 Methods . . . 166
7.2.1 Apparatus and Stimuli . . . 166

7.2.2 Procedure . . . 166
7.2.3 Observers . . . 167
7.3 Results . . . 169
7.3.1 Monocular data . . . 171
7.3.2 Dichoptic data . . . 173
7.4 Binocular summation at threshold . . . 173
7.5 Models of amblyopia . . . 177
7.5.1 Independent channels . . . 178
7.5.2 Weight of interocular suppression . . . 180
7.5.3 Attenuator models . . . 182
7.5.4 Varying the saturation constant . . . 183
7.5.5 Stochastic models . . . 185
7.6 Discussion . . . 188
7.6.1 Comparison with previous studies . . . 188
7.6.2 Importance of neural binocular summation . . . 189
7.6.3 Treatment of amblyopia . . . 190
7.6.4 Origins of the noise . . . 191
7.6.5 Conclusions . . . 192

8 Summary and Discussion **193**

8.1 Summary of findings . . . 193
8.2 Model development . . . 196
8.2.1 Interpretation of parameter values . . . 198
8.2.2 Neurophysiological loci of model components . . . 199
8.2.3 Assumptions of the model, and alternative approaches . . . 200
8.3 Common trends across experiments . . . 201
8.3.1 Binocular summation across experiments . . . 201
8.3.2 Individual differences in dichoptic masking . . . 202

8.4 Relation to other work . . . 203
8.4.1 Classic psychophysics . . . 203
8.4.2 Neurophysiology . . . 205
8.4.3 Ding & Sperling . . . 205
8.4.4 Rivalry models . . . 206
8.5 Future directions . . . 208
8.5.1 Binocular summation: bandwidth and luminance dependence . . . 208
8.5.2 Masking effects in other stimulus domains . . . 208
8.5.3 Cross-channel masking in amblyopia . . . 210
8.6 Conclusions . . . 210

References **212**

Appendix **226**

A Stereoscope calibration **227**

A.1 Introduction . . . 227
A.2 Design and theory . . . 227
A.3 Calibration procedure . . . 228

List of Figures

1.1 Example dipper function for contrast discrimination, and sigmoidal transducer. 21

1.2 Diagram of Legge & Rubin's (1981) contrast matching paradigm. 29

1.3 Summary of chapter layout and interrelationships. 33

1.4 Summary of thesis as a map of stimulus space. 34

2.1 Example of a horizontal sinusoidal grating with a raised cosine spatial window. 39

2.2 Schematic diagrams of the mirror stereoscope. 42

2.3 Example probit fit to psychometric data. 45

3.1 Monocular, binocular and dichoptic contrast discrimination functions. 48

3.2 Psychometric slope data for Experiment 1. 49

3.3 Schematic diagrams of four model architectures. 51

3.4 Best fits of four models to the data of Experiment 1. 52

3.5 Psychometric slope predictions for Experiment 1. 58

3.6 Model response surfaces on linear and logarithmic axes. 59

3.7 Results and model predictions for Experiment 2. 61

3.8 Results of an extended version of Experiment 2 (discrimination surface). . . . 62

3.9 Contrast matching data for two subjects with Two Stage model predictions. . 65

3.10 Predictions of the Twin Summation model for matching data. 66

3.11 Psychometric surface for dichoptic masking. 68

3.12 Paradoxical dichoptic psychometric functions for three observers. 70

3.13 Response surfaces for the Two Stage and Twin Summation models. 71

4.1 Diagram of stimulus configurations for cross-channel experiments. 74
4.2 Interocular cross-channel masking results using an orthogonal mask stimulus. 76
4.3 Interocular cross-channel masking results using an oblique mask stimulus. . . 77
4.4 Binocular summation assessed across the entire masking function. 79
4.5 Cross-channel masking at varying stimulus durations. 81
4.6 Cross-channel masking at varying stimulus durations, normalized to baseline. 81
4.7 Illustration of possible model configurations, with canonical predictions. . . . 83
4.8 Best fit of Model 1 to the monocular and dichoptic data of Experiment 5. . . 85
4.9 Predictions of Model 1 for remaining conditions. 86
4.10 Half binocular model and predictions. 87
4.11 Best fits of Models 1 - 3 to the data of subject TAY. 88
4.12 Temporal model fits to the data of Experiment 6. 91

5.1 Results of the cross-channel twin mask experiment. 98
5.2 Model diagram showing where cross-channel suppression can impact on the two stage model. 100
5.3 Model predictions for three locations of monocular cross-channel masking. . . 101
5.4 RMS errors for nine model configurations, fit to the data of Experiment 7. . . 102
5.5 Temporal sequence for each trial of the adaptation experiment. 105
5.6 Results of the adaptation experiment. 106
5.7 Model fits to adaptation data. 109
5.8 Diagram of the feedback model. 114
5.9 Channel activation in the feedback model. 115
5.10 Behaviour of the feedback model. 115
5.11 Stimuli and temporal waveforms for Experiment 9. 117
5.12 Spatial frequency tuning of orthogonal masking. 119
5.13 Temporal frequency tuning of orthogonal masking. 119
5.14 Stimuli and temporal waveforms used in Experiment 10. 122

5.15 Monocular and dichoptic cross channel masking at different spatiotemporal frequencies. 123

5.16 Threshold normalized masking functions, collapsed across spatiotemporal condition. 125

5.17 Diagram of cross-channel masking model. 127

5.18 Change in RMS error with alpha. 128

5.19 Model fits to normalized data. 129

5.20 Best fitting model weights for monocular and dichoptic conditions, plotted against speed. 130

5.21 Best fits of two models to the threshold-normalized dichoptic data of Experiment 10. 132

5.22 Dichoptic data from Chapter 4, normalized to detection threshold. 133

6.1 Antiphase dichoptic masking functions 138

6.2 Diagram dichoptic phase model. 140

6.3 Phase tuning of dichoptic masking 141

6.4 Spatial frequency tuning functions replotted from Legge (1979). 142

6.5 Dichoptic spatial frequency tuning functions for in-phase and antiphase gratings. 144

6.6 Dichoptic orientation tuning functions for in-phase and antiphase gratings. . 146

6.7 Further dichoptic orientation tuning functions. 148

6.8 Best fits of the orientation model. 150

6.9 Activation in the orientation model. 151

6.10 Best fits of the spatial frequency model. 154

6.11 Monocular spatial frequency tuning functions 156

6.12 Monocular orientation tuning functions 157

6.13 Comparison of monocular tuning data and the dichoptic model. 158

6.14 Illustration of stimuli used to measure binocular fusion. 160

6.15 Dichoptic masking surfaces. 162

6.16 Predictions of the model for orthogonal masks. 163

7.1 Contrast discrimination functions for normal control subject. 170

7.2 Monocular and binocular contrast discrimination functions for amblyopic observers. 171

7.3 Dichoptic and binocular contrast discrimination functions for amblyopic observers. 172

7.4 Summary of monocular facilitation. 173

7.5 Normalized dichoptic masking data for all amblyopes. 174

7.6 Binocular summation ratios for three normal observers. 175

7.7 Binocular summation ratios for eight amblyopic observers (standard method). 175

7.8 Binocular summation ratios for eight amblyopic observers (normalization method). 177

7.9 Amblyopic and normal contrast masking functions. 178

7.10 Behaviour and diagram of the independent channels model. 179

7.11 Dichoptic masking functions with different interocular weights. 181

7.12 Best fits of the early attenuator model. 183

7.13 Best fits of the late attenuator model. 184

7.14 Model fits with independent saturation constants. 185

7.15 Model fits with independent saturation constants and an attenuator. 186

7.16 Behaviour and diagram of the stochastic model. 187

8.1 Diagram of full two stage model. 197

8.2 Binocular summation ratios from three experiments. 201

8.3 Dichoptic masking summarised across several experiments and observers. . . 203

8.4 Comparison of the Ding & Sperling and Two Stage models. 206

8.5 Schematic diagrams of three binocular rivalry models. 207

A.1 Schematic diagram of the stereoscope set up. 228

A.2 Image developed for use in the calibration procedure. 230

List of Tables

1.1 Meta-analysis of binocular summation ratios. 27

3.1 Parameters and error statistics for Experiment 1. 53

3.2 Model parameters and RMS errors for Experiment 2. 63

4.1 Best fitting parameters and error statistics for Model 1. 85

4.2 Best fitting parameters for the temporal model. 91

5.1 Configurations of monocular and dichoptic cross-channel masking. 99

5.2 Parameters and error statistics for the adaptation model. 109

5.3 Parameters and error statistics of two models for the dichoptic data of Experiment 10. 132

6.1 Parameters and RMS errors of the orientation model. 149

6.2 Parameters and RMS errors of the spatial frequency model. 154

7.1 Demographic and clinical details of amblyopic subjects. 168

7.2 Dichoptic masking slopes for all data sets. 174

7.3 Parameters and error statistics for the attenuator models. 183

7.4 Parameters and error statistics for independent saturation constant models. . 184

Chapter 1

Introduction

1.1 Binocular vision

A common feature of all primates, and many other species such as cats, dogs, and birds of prey, is that the two eyes are arranged to have largely overlapping visual fields. This configuration confers distinct advantages, most notably depth cues from stereopsis, and a degree of redundancy if one eye is damaged. However, the brain is left with a problem; it must combine two similar images into a single binocular percept, from which it can extract meaningful information about the world.

The processes by which binocular combination occurs are some of the most extensively studied in the field of vision research (see Howard, 2002). Early work focussed on binocular perception of brightness when varying luminances are presented to the two eyes (Sherrington, 1904; Levelt, 1965a). However, absolute luminance information is largely ignored by the visual system in order to allow it to operate over a wide range of light levels, from near darkness to bright sunlight. This is possible, in part, because of a sophisticated process of retinal gain control mechanisms, which produce a normalised output to compensate for changes in the absolute level of light entering the eye.

One consequence of this is that relatively little information about object luminances in a visual scene ever reaches the brain. Instead, visual cortex receives information about the *relative* luminances, also termed the image contrast. Neurophysiological studies of primary visual cortex in cat have confirmed the presence of cells sensitive to luminance modulations, such as bar or grating stimuli (Hubel and Wiesel, 1959). For this reason, much visual research over the last 50 years has focussed on contrast perception, and the study of binocular vision is no exception.

This thesis is concerned with human contrast perception, and specifically the processes

by which the brain combines contrast signals presented to the two eyes. The introduction begins with a review of some important findings from the literature on contrast vision, and continues with a summary of previous work focussing on binocular combination of signals in both normal and amblyopic humans. The chapter concludes by summarising the remainder of the thesis.

1.2 Important findings in contrast perception

1.2.1 Detection thresholds and contrast sensitivity

Our sensitivity to the contrast of a periodic stimulus, such as a sinusoidal luminance variation (grating), is dependent on that grating's spatial and temporal properties. Spatial frequency is measured in cycles per degree (cpd) of visual angle, with low frequencies corresponding to few cycles, and high frequencies to many cycles, per unit area.

We can measure an observer's ability to detect stimuli of different spatial (or temporal) frequencies using psychophysical methodology. Small changes in stimulus magnitude are related to subject performance over many trials. The contrast sensitivity function (CSF), which plots sensitivity against spatial frequency (Schade, 1956), for normal humans approximates an inverted 'u' shape, peaking at around 3-5cpd (for a fixed size of stimulus), and falling off at lower and higher frequencies (Campbell and Robson, 1968; see chapter 5 of DeValois and DeValois, 1988, for a review).

It is generally believed that the shape of the CSF is the envelope of a number of bandpass detecting mechanisms, each of which is preferentially sensitive to a limited range of spatial frequencies. Evidence for this multiple mechanism hypothesis comes from psychophysical studies using techniques such as adaptation (Blakemore and Campbell, 1969; Georgeson and Harris, 1984) or subthreshold summation (Graham and Nachmias, 1971; Graham, 1989), and also from neurophysiological investigation of the tuning properties of neurones in visual cortex (Campbell, Cleland, Cooper and Enroth-Cugell, 1968; Campbell, Cooper and Enroth-Cugell, 1969; Maffei and Fiorentini, 1973; Movshon, Thompson and Tolhurst, 1978; Tolhurst and Thompson, 1981; De Valois, Albrecht and Thorell, 1982a; De Valois, Yund and Hepler, 1982b). It is likely that these mechanisms, often termed channels, are the beginnings of our analysis of visual scenes, which forms the basis of our perceptual experience of the world (Marr, 1983).

1.2.2 Within-channel masking

Nachmias and Sansbury (1974) found that detection of a test grating can be improved by the

addition of a low contrast background grating with the same spatial properties as the test (similar findings had also been reported previously, i.e. Fig. 2 of Campbell and Kulikowski, 1966). This additional grating is known as a pedestal, and the improvement in performance conferred by it is termed facilitation. A simplistic analogy for this effect is that of filling a glass with water. If threshold is reached when the glass overflows, beginning with it half-full will require less water to be added than beginning with it empty[1].

As pedestal contrast is increased, the amount of facilitation also increases, until pedestal contrast is just above the detection threshold (the reciprocal of sensitivity) for the test grating alone. After this point, when the pedestal itself is visible, detection thresholds begin to increase again, in proportion to the pedestal contrast. In this part of the contrast range, the pedestal is said to be 'masking' the test, as more test contrast is required to reach threshold than if the pedestal were absent. In addition, the subject is now detecting a contrast *increment* on a visible pedestal, hence the task is contrast discrimination.

When threshold is plotted against pedestal contrast (known as a TvC, or threshold versus contrast, plot), these two effects of the pedestal - masking and facilitation - produce a 'dipper' shaped function (see Fig. 1.1A). Dipper functions have been studied extensively over the last few decades (i.e. Campbell and Kulikowski, 1966; Tolhurst and Barfield, 1978; Legge and Foley, 1980; Foley, 1994; Ross and Speed, 1996; Bird, Henning and Wichmann, 2002) and are a robust and pervasive phenomenon, replicated numerous times in humans, and even recorded behaviourally in monkey (Smith III, Harweth, Levi and Boltz, 1982). They occur at all orientations and spatial frequencies, and persist in the presence of additional masking gratings with different spatial properties (Foley, 1994; Ross and Speed, 1996; Holmes and Meese, 2004), or broadband noise (Legge, Kersten and Burgess, 1987; Henning and Wichmann, 2007). Masking, and some facilitation, is also preserved if the pedestal spatial frequency is altered slightly, relative to that of the test spatial frequency (Legge and Foley, 1980; Tolhurst and Barfield, 1978). Several explanations have been offered for their existence, which are discussed in the following section.

1.2.3 Modelling the dipper function

An early account of the dipper function was offered by Legge and Foley (1980), who proposed a model which featured a saturating transducer function (see also Foley and Legge, 1981). Figure 1.1B shows an example transducer function, generated using the equation,

[1]Contemporary explanations of the effect are more complex than this, and generally rely on signal detection theory (Green and Swets, 1966) rather than the high threshold assumptions of the glass of water analogy (see Nachmias, 1981; Tyler and Chen, 2000). However, it remains an intuitive aid to the reader's understanding.

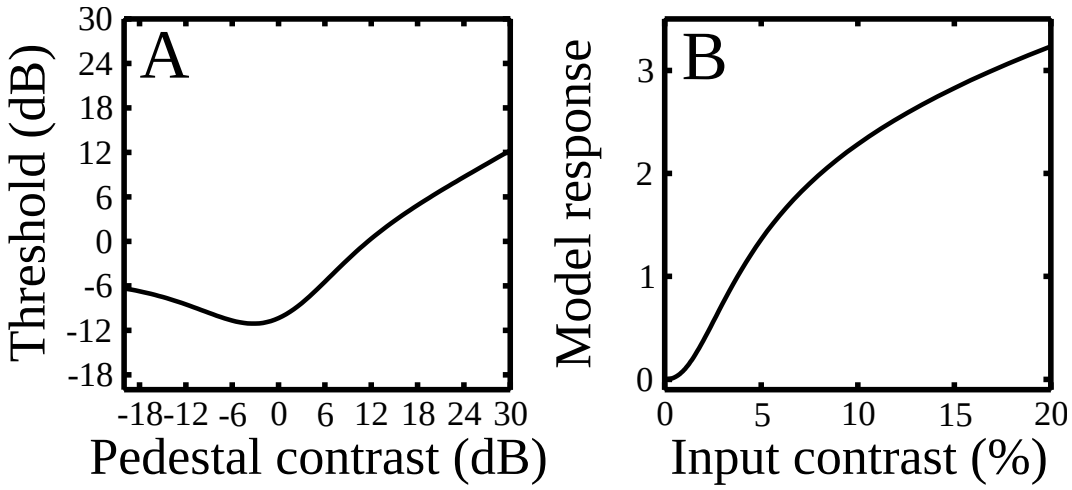

Figure 1.1: A) Example dipper function for contrast discrimination. Abscissa gives the contrast of a background pedestal, with the same spatial properties as the test stimulus. Ordinate gives the contrast required for reliable detection of the test increment. At low pedestal contrasts, thresholds decrease (facilitation), before increasing linearly at higher pedestal contrasts (masking). Contrast values are given in decibels (dB), which are logarithmic units, as described in Chapter 2. B) Sigmoidal transducer function proposed by Legge and Foley (1980) to explain the dipper function, plotted on linear axes. See text for details.

$$resp = \frac{C^p}{Z + C^q}, \tag{1.1}$$

where C is input contrast, and p, q and Z are model parameters. A sigmoidal (s-shaped) transducer function such as this can produce the dipper function for contrast discrimination by taking the derivative of the curve. This means that thresholds depend on the gradient of the transducer - the increasing gradient in the S-shaped response curve at low input contrasts gives an increasing amount of facilitations, while the decreasing gradient at high contrasts produces a progressive masking effect as contrast increases.

Interestingly, the equation used by Legge and Foley (1980) is similar to the Naka-Rushton function (Naka and Rushton, 1966) frequently used (e.g. Heeger, 1992) to describe the contrast response of individual V1 neurones (as well as many other neurones). The main difference is that in the Naka-Rushton equation, the exponents p and q typically take on the same value, meaning that the response function saturates, restricting the neurone's dynamic range. However, recent models have found that pooling populations of such neurones with a range of parameter values (i.e. sensitivities) can produce an overall response function which closely resembles the Legge and Foley (1980) transducer function, and hence can predict the dipper function for contrast discrimination (Watson and Solomon, 1997; Chirimuuta and Tolhurst, 2005b). Furthermore, there is fMRI evidence that the response of V1 to grating contrast closely follows the sigmoidal transducer (Boynton, Demb, Glover and Heeger, 1999), indicating that V1 may be the site at which contrast discrimination performance is

determined.

An alternative explanation to the nonlinear transducer function is that dipper functions arise because of two separate process - uncertainty reduction and multiplicative noise. Pelli (1985) developed a model (also considered by Foley and Legge, 1981) which assumes that performance is limited by uncertainty about which channels should be monitored (see next section for a discussion of channels) to determine an observer's response. In this framework, a pedestal reduces uncertainty by increasing the signal-to-noise ratio in the relevant channels, allowing the irrelevant channels to be ignored, and thus improving performance. At higher pedestal contrasts, performance is limited by the variance of the signal, which is assumed to increase in proportion to the signal strength (i.e. multiplicative noise; Kontsevich, Chen and Tyler, 2002; McIlhagga and Peterson, 2006; Burton, 1981). This produces the handle portion of the dipper function and has strong physiological plausibility, given that neurones are known to have multiplicative noise (Tolhurst, Movshon and Thompson, 1981b).

These two explanations are the best established, however they are not the only ones. Recently, stochastic resonance - the improvement in the information capacity caused by noise in a system (i.e. Wiesenfeld and Moss, 1995) - has been used to explain facilitation (Blackwell, 1998; Sasaki, Todorokihara, Ishida, Miyachi, Kitamura and Aoki, 2006). Furthermore, it is not clear that the static transducer and the uncertainty reduction/multiplicative noise account are mutually exclusive. Instead, they may simply be different levels of explanation of the same phenomenon.

1.2.4 Multiple channels in human vision

A central concept in the psychophysical investigation of vision is that the visual system comprises distinct mechanisms, or channels, which are selectively sensitive to stimuli of specific orientations and spatiotemporal frequencies (Blakemore and Campbell, 1969). Evidence for this arrangement comes from a variety of methods, including adaptation (Blakemore and Campbell, 1969; Blakemore and Nachmias, 1971; Georgeson and Harris, 1984), subthreshold summation (Kulikowski, Abadi and King-Smith, 1973; Graham and Nachmias, 1971) and contrast masking (Campbell and Kulikowski, 1966; Pelli, 1981). A detailed comparison of the various methods is given by Graham (1989). Recently, Watson (2000) found that a model featuring multiple channels with bandwidths of 1.4 octaves of spatial frequency spaced 45° apart provided a better fit to detection thresholds for a range of stimuli (the Modelfest data) than four alternative models (see also Watson and Ahumada, 2005).

Historically, there have been substantial differences in estimates of channel bandwidths, depending on the experimental method used. For example, adaptation studies have produced bandwidth estimates of around 7° for orientation (Blakemore and Nachmias, 1971) and 1-1.4

octaves (an octave is a change in frequency by a factor of 2) of spatial frequency (Blakemore and Campbell, 1969; Georgeson and Harris, 1984). Using a masking paradigm, Campbell and Kulikowski (1966) estimated the orientation bandwidth to be between 12 and 15° whereas Kulikowski *et al.* (1973) found bandwidths as narrow as 3° using subthreshold summation.

The narrow bandwidths reported by early studies may in part be due to the use of spatially extensive stimuli. Probability summation over area for large stimuli can lead to artificially narrow bandwidth estimates because of phase cancellation between stimulus components in some regions of the image (Stromeyer and Klein, 1975; Wilson and Bergen, 1979; Bergen, Wilson and Cowan, 1979). Studies which have taken this into account have produced broader bandwidth estimates of around 1.75 octaves (Bergen *et al.*, 1979) or 1.25-2.5 octaves (Wilson, McFarlane and Phillips, 1983) of spatial frequency, and 15-30° of orientation (Phillips and Wilson, 1984). These broader bandwidth estimates compare favourably with findings from single cell neurophysiology (Campbell and Robson, 1968; Campbell *et al.*, 1969; Maffei and Fiorentini, 1973; Ikeda and Wright, 1975; Movshon *et al.*, 1978; Tolhurst and Thompson, 1981; De Valois *et al.*, 1982a,b).

The contrast discrimination dipper functions discussed in previous sections have come to be understood as a within-channel phenomenon, as pedestal and test activate the same detecting mechanism. They can, therefore, tell us much about the characteristics of an individual mechanism, but little about how groups of such mechanisms interact with each other. In the following section, experiments are discussed which explore interactions between channels.

1.2.5 Cross-channel masking

When a mask falls outside the bandwidth of the mechanism which detects the test, it produces no excitation in the test mechanism. However, our ability to detect a test grating can still be affected by such a mask (Campbell and Kulikowski, 1966; Phillips and Wilson, 1984). This phenomenon is often called cross-orientation masking, as many studies, both psychophysical and neurophysiological, have used orthogonal gratings as stimuli (Morrone, Burr and Maffei, 1982; Nelson, 1991; Meese and Holmes, 2007). Here a broader term, cross-channel masking, is used to acknowledge the fact that the masking stimuli can differ in either the orientation or spatial frequency domains (or both) in order to meet the requirement of activating a separate detecting mechanism from that of the test (Foley, 1994; Meese and Holmes, 2002).

Cross-channel masks raise detection thresholds in proportion to mask contrast, producing monotonically increasing masking functions (i.e. Meese and Holmes, 2002). Foley (1994) showed that a cross-channel mask will displace a dipper function upwards and to the right,

such that the dip region is preserved and the dipper handles superimpose at high pedestal contrasts. This is an important finding, as it suggests that the mask is processed independently of the pedestal and test stimuli. In a similar vein, the substantial facilitation (dip) observed using within-channel masks does not occur for cross-channel stimuli (Foley, 1994), although weak facilitation is sometimes observed (Meese and Holmes, 2007).

It has generally been supposed that cross-channel masking is caused by some suppressive interaction between cortical neurones, perhaps as a means of sharpening their orientation and spatial frequency tuning (Morrone *et al.*, 1982; Heeger, 1992). However, a number of recent findings in the physiology literature have cast doubt on this assertion. Freeman, Durand, Kiper and Carandini (2002) showed that suppression still occurred with mask stimuli flickering too fast to excite cortical neurones. Indeed, the pattern of spatial and temporal tuning is more consistent with cells in the lateral geniculate nucleus (LGN) than visual cortex (Priebe and Ferster, 2006; Li, Peterson, Thompson, Duong and Freeman, 2005; Li, Thompson, Duong, Peterson and Freeman, 2006). These findings are discussed in depth in Chapter 5.

1.3 Contrast gain control

There is much evidence that the visual system utilises population responses from large numbers of neurones in order to produce a unique neural representation of different stimuli (see Georgeson, 2004). However, since neurones saturate at high contrasts (i.e. Geisler and Albrecht, 1992) two different high contrast stimuli could in principle produce the same pattern of activation, despite having very different properties. To overcome this difficulty, the visual system must keep neurones within their optimal operating range, so that only those being optimally stimulated produce strong responses. This is achieved by a divisive 'gain pool' (Heeger, 1992), in which neurones reduce each others' firing by mutual inhibition.

This process, termed contrast gain control, has been used to explain both single-cell and psychophysical behaviour (i.e. Ohzawa, Sclar and Freeman, 1982, 1985; Heeger, 1992; Foley, 1994). A gain control operates like an automatic volume control - weak inputs are amplified, and strong inputs are attenuated. This may allow neurones to remain within their optimal operating range (Tolhurst, Dean and Thompson, 1981a; Albrecht and Hamilton, 1982) across the wide variation of contrasts found in natural images (Lauritzen and Tolhurst, 2005), and after prolonged exposure (adaptation) to high contrast images (Ohzawa *et al.*, 1985; Maattanen and Koenderink, 1991; Wilson and Humanski, 1993).

In neurophysiology, gain control models have been shown to explain a variety of data from cortical simple cells better than previous models (i.e. Tolhurst and Dean, 1987), which

relied on a threshold nonlinearity (Tolhurst and Heeger, 1997). Psychophysical models incorporating a divisive gain control can produce both within-channel dipper functions and cross-channel masking effects (Foley, 1994), and are now widespread (i.e. Watson and Solomon, 1997; Meese and Holmes, 2002; Olzak and Thomas, 2003; Holmes and Meese, 2004; Ding and Sperling, 2006a,b). Recent models have met with some success in describing the responses of neural populations to natural images (Clatworthy, Chirimuuta, Lauritzen and Tolhurst, 2003) as well as reproducing psychophysical discrimination (Geisler and Albrecht, 1997; Chirimuuta and Tolhurst, 2005b) and identification (Chirimuuta and Tolhurst, 2005a) behaviour. Gain control has proved useful in numerous other experimental paradigms, such as edge detection (Hammett, Georgeson, Bedingham and Barbieri-Hesse, 2003), direction selectivity (Albrecht and Geisler, 1991), and suppression in the LGN (Bonin, Mante and Carandini, 2005) to list but a few examples. The ubiquity of gain control processes throughout the visual system indicates their importance for efficient visual coding.

Despite the widespread interest in contrast gain control, there are still aspects of it which are poorly understood. One aim of this thesis is to characterise the suppressive gain control interactions which occur within and between the eyes (see next section). It should be stressed that the intention is to investigate the underlying mechanisms rather than try to explain their purpose within the broader context of vision.

1.4 Presenting stimuli to different eyes

The studies described above all used stimuli which were viewed binocularly, with both eyes seeing the same image. By varying the stimuli presented to each eye, in either their spatiotemporal properties or their contrasts, inferences can be drawn about the neural processes which combine information from the two eyes into a single cyclopean image.

From neuroanatomical work, it has been known for some time that monocular signals remain separate until they reach primary visual cortex (also known as V1, or striate cortex). Hubel and Wiesel (1959, 1962) found that neurones in V1 have differing degrees of binocularity. Some cells can be excited equally by both eyes, whilst others have a greater response to one eye over the other (ocular dominance). Psychophysically, binocular combination can be considered as a unitary process, which presumably pools activation over large populations of cells with differing degrees of binocularity, or considers only the final (fully binocular) output stage (see Anderson and Movshon, 1989). There is evidence that humans do not have access to purely monocular neurones, as they are unable to perform utrocular discrimination - identifying to which eye a stimulus has been presented (Blake and Cormack, 1979; Barbeito, Levi, Klein, Loshin and Ono, 1985).

1.4.1 Binocular summation at threshold

In their classic study on binocular summation, Campbell and Green (1965) found that the ratio of binocular to monocular sensitivities in a grating detection task was around 1.4 ($\sqrt{2}$). This is greater than the advantage expected from probability summation of the two inputs (∼1.25 or less; Pirenne, 1943; Eriksen, 1966) but less than if the signals were summed linearly (a ratio of 2, assuming late noise). As Campbell and Green (1965) point out, $\sqrt{2}$ is the improvement expected from summing two uncorrelated noisy detectors. However, there is a problem with this interpretation, as it assumes that under monocular stimulation, noise from the unstimulated eye can be discounted (Blake, Sloane and Fox, 1981b; Campbell and Green, 1965), which seems unlikely. An alternative explanation is that signals from the two eyes are subject to some nonlinear neural processing before they are summed, and numerous models have been proposed to account for this (Legge, 1984b; Anderson and Movshon, 1989; Meese, Georgeson and Baker, 2006).

One of the most influential explanations was proposed by Gordon Legge (1984b). Legge suggested that the inputs to the two eyes are squared before being summed together, with the output being the square root of the summed value:

$$B = \sqrt{L^2 + R^2}, \tag{1.2}$$

where L and R are the input contrasts to the left and right eyes, and B is the binocular response. This scheme, termed quadratic summation, gives the required ratio of $\sqrt{2}$ (to verify this, replace L and R with 0 and 1, or 1 and 1).

Although summation ratios of $\sim\sqrt{2}$ have been recorded in numerous studies over the intervening decades since Cambell and Green's paper, several investigators have reported values which differ greatly from this. For example, Snowden and Hammett (1996) report one observer with a ratio of 1.73, and Simmons (2005) found a mean ratio across four observers of around 1.6. Using a reaction time paradigm, Harwerth, Smith III and Levi (1980) report ratios at near-threshold contrasts of between 1.44 and 1.74 for eight subjects. Georgeson and Meese (2005) performed a systematic study of binocular summation at threshold, using stimuli over a range of temporal frequencies and found that summation decreased slightly with speed (TF/SF), and had a mean ratio of around 1.7.

A detailed summary of binocular summation ratios in the literature is given in Table 1.1. The table only includes studies which used sinusoidal gratings as stimuli, however similar findings have also been obtained using flashes of light (Westendorf and Fox, 1974), as well as other psychophysical tasks (see Blake and Fox, 1973 and Blake *et al.*, 1981b for reviews of early work).

Source	Mean summation ratio	Range of summation ratios	Spatial Frequencies (cpd)	Temporal properties	No of subjects	Method	Setup
Campbell & Green (1965)	1.42	1.37 - 1.44	2 - 45	Constant	2	MOA	Occluder
Lema & Blake (1977)	1.35	1.29 - 1.50	2 - 6.5	1.5Hz/Constant	4	MOA	Occluder
Blake & Levinson (1977)$^{\epsilon}$	1.42	1.23 - 1.60	1 - 9.5	Constant	1	MOA	Stereoscope
Derefeldt *et al.* (1979)	1.47	0.75 - 2.04	0.5 - 40	Constant	33	MOA	Patch
Rose (1978)	1.48	0.83 - 2.37	0.5 - 30	3.5Hz/Constant	8	MOA	Occluder
Birch (1979)†	-	1.24 - 1.83	0.25 - 1	N/S	1	Matching	Unknown
Blake & Rush (1980)$^{\epsilon}$	1.42	1.32 - 1.48	0.5 - 8	3.5Hz/250ms	3	MOA	Stereoscope
Blake & Rush (1980)$^{\epsilon}$	1.47	1.39 - 1.66	1	4Hz/1000ms	2	2AFC	Stereoscope
Harwerth *et al.* (1980)	-	1.44 - 1.74	4	500ms	8	RT	Patch/diffuser
Levi *et al.* (1980)$^{\epsilon}$	1.44	1.31 - 1.58	0.5 - 8	N/S	1	2AFC	Stereoscope
Rose (1980)$^{\epsilon}$	1.59	0.85 - 2.41	0.5 & 5	2 - 30Hz	6	MOA	Occluder
Arditi *et al.* (1981)$^{\epsilon}$	1.63	1.28 - 1.71	0.3 - 9.6	0.5 - 16Hz	1	2AFC	Stereoscope
Legge (1984a)	1.55	1.50 - 1.64	0.5	200ms	4	2AFC	Split
Gilchrist & McIver (1985)$^{\epsilon}$	1.39	1.03 - 1.86	0.75 - 18	Constant	2*	MOA	N/S
Ross *et al.* (1985)$^{\epsilon}$	1.08	1.02 - 1.14	0.4 - 19.23	N/S	70	2AFC	Patch
Holopigian *et al.* (1986)$^{\epsilon}$	1.28	1.10 - 1.57	0.5 - 8	500ms	2	2AFC	Stereoscope
Gilchrist & Pardhan (1987)	1.47	1.17 - 2.16	6	N/S	8	2AFC	N/S
Rose *et al.* (1988)$^{\epsilon}$	1.54	1.42 - 1.61	0.75 - 6	60ms	3	2AFC	Stereoscope
Anderson & Movshon (1989)$^{\epsilon}$	1.41	1.28 - 1.48	2.6 - 5	Constant	3	MOA	Stereoscope
Denny *et al.* (1991)$^{\epsilon}$	1.65	1.21 - 2.26	2 - 20	500ms	3	2AFC	Stereoscope?
Pardhan (1996)	1.47	1.46 - 1.48	1 & 6	Constant	8	2AFC	Patch
Snowden & Hammett (1996)	1.51	1.22 - 1.77	6	2Hz/900ms	3	2AFC	Goggles
Simmons & Kingdom (1998)$^{\epsilon}$	1.63	1.57 - 1.69	0.5	200ms	2	2AFC	Goggles
Pardhan & Rose (1999)$^{\epsilon}$	1.47	-	4	200ms	4	2AFC	Stereoscope
Marshman *et al.* (2001)$^{\epsilon}$	1.33	1.24 - 1.41	0.5 - 17.8	5.5s	14	2AFC	Occluder
Hood & Morrison (2002)$^{\epsilon}$	1.25	1.16 - 1.33	10 - 40	Constant	9	MAL	Occluder
Gagnon & Kline (2003)$^{\epsilon}$	1.69	1.21 - 2.75	1 - 18	N/S	28	3AFC	Patch
Valberg & Fosse (2002)$^{\epsilon}$	1.85	1.35 - 2.00	0.3 - 25	Constant	10	MOA	N/S
Meese & Hess (2004)	1.96	1.65 - 2.55	0.46 - 1	200ms	2	2AFC	Stereoscope
Meese & Hess (2005)	1.81	1.62 - 2.01	0.46	200ms	2	2AFC	Stereoscope
Maehara & Goryo (2005)$^{\epsilon}$	1.47	1.28 - 1.73	2	33ms	3	2AFC	Stereoscope
Simmons (2005)$^{\epsilon}$	1.63	1.61 - 1.65	0.5	200ms	4	2AFC	Stereoscope
Meese *et al.* (2006) Expt 1	1.70	1.53 - 1.78	1	40 - 500ms	2	2AFC	Goggles
Medina & Mullen (2007)	1.70	1.18 - 2.22	1.5	2 - 16Hz	3	2AFC	Patch
Present thesis	1.65	1.48 - 1.83	0.5 - 3	50 &200ms	7	2AFC	Both
Overall	**1.52**	**0.75 - 2.75**					

Table 1.1: Meta-analysis of binocular summation ratios at and around detection threshold. Ratios refer to bin/mon sensitivities for in-phase achromatic sinusoidal gratings and are rounded to 2 decimal places. Temporal information is given as duration for static stimuli, and as temporal frequency for flickering or moving stimuli. For papers containing data on clinical populations (i.e. amblyopia, AMD), only the results for normal (control) subjects have been included. MOA = Method of Adjustment, nAFC = n alternative forced choice, RT = reaction time, MAL = Method of Ascending Limits, Split = split screen viewing, N/S = not stated, *abnormal observers removed, †reported in Cogan (1987), $^{\epsilon}$some data estimated from figures. 'Both' indicates the use of both goggles and stereoscope.

The range of summation ratios in Table 1.1 is broad. Mean values generally fall between $\sqrt{2}$ and 2, with some outliers. To some extent, these differences may be attributable to methodological differences in psychophysical task and stimulus properties. Another factor which could influence results is the fate of the unstimulated eye during monocular presentation. In many early studies, an eye patch or occluder was used, which creates a luminance imbalance between the eyes. Later studies favoured stereoscopes or shutter goggles to ensure that equal luminance was presented to both eyes. From the studies in Table 1.1, the mean summation ratio with a patch or occluder is 1.44 (n=11), but when the eyes viewed equal luminance the mean is 1.55 (n = 17). However, the two presentation methods are yet to be compared in a single study, using the same observers.

The weight of evidence suggests that the classic $\sqrt{2}$ improvement of two eyes over one is by no means a fixed constant. This, in turn, casts doubt on the validity of models (such as Legge's) which rely on this relationship for both threshold and suprathreshold behaviour.

1.4.2 Binocular summation above threshold

For contrast discrimination tasks, one might expect that the binocular advantage seen at threshold would extend across the whole dipper function. However, Legge (1984a) found that the 'handles' of monocular and binocular dippers converged at high pedestal contrasts. One interpretation of this finding is that binocular summation is lost at higher contrast levels. Legge (1984b) did not conclude this, because his quadratic summation model predicted this behaviour. According to the model, the monocular dipper is shifted both upwards and to the right by the same factor ($\sqrt{2}$), causing the linear 'handle' portions to superimpose. Binocular summation is not lost, it is simply hidden at higher contrasts.

As Meese *et al.* (2006) point out, the monocular-binocular dipper comparison contains an elementary confound. Relative to the monocular condition, the binocular condition has both an extra test signal and an extra pedestal. A more appropriate manipulation is to compare performance in the binocular condition to performance when the test is presented only to one eye, but the pedestal is still presented to both. This configuration (termed the half-binocular condition) reveals that binocular summation is preserved across the entire contrast range, and has a constant ratio of around 1.8.

1.4.3 Interocular contrast matching

Detection and discrimination are performance based measures, which tell us about an observer's ability to carry out a task. However, they tell us very little about the subjective experience of viewing stimuli. Perceptual measures, such as contrast matching, allow us to

explore how presenting different contrast levels to each eye actually appears to the observer.

Legge and Rubin (1981) performed an important experiment, in which subjects matched gratings with binocularly unequal contrasts, to a binocularly equal standard[2]. The unequal test gratings were held at a fixed ratio, so that the response could lie on a line radiating from the origin, when plotted as left eye against right eye contrast (dotted lines in Figure 1.2). The results over a range of different ratios defined a contour of perceptually equivalent binocular contrast pairs (curves in Figure 1.2). At low standard contrasts, these contours were relatively linear (grey function). However, as standard contrast increased, the contours became more curved, and doubled back on themselves (black function). These results show that as contrast in one eye increases from zero, it is necessary to increase contrast in the other eye also, in order to achieve a match. This is the opposite of what we might expect, were the visual system simply summing the two inputs together.

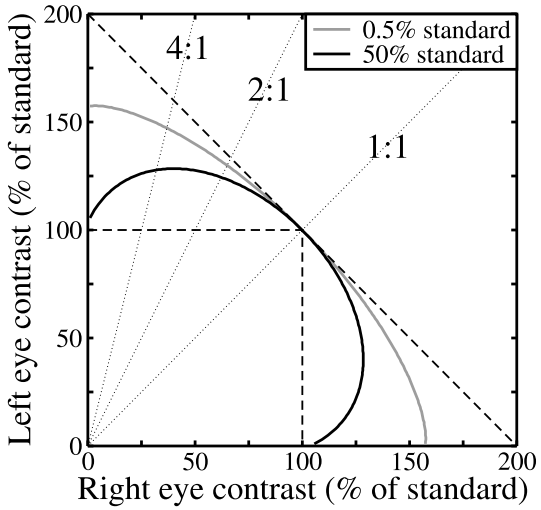

Figure 1.2: Diagram of the Legge and Rubin (1981) contrast matching paradigm. Perceptual matches to a binocularly equal standard (here 0.5% or 50%) were made using binocularly unequal matching stimuli. The contrasts of the matching stimuli were varied in ratio, given by the dotted diagonal lines radiating from the origin, which illustrate three possible ratios. The point at which standard and match appeared equal depended on the ratio, and produced curves like those shown. The curves fell in between the oblique dashed line, which indicates perfect linear summation of inputs, and the dashed square, which indicates no summation.

This unusual result is very similar to a well known effect found in the brightness matching literature, called Fechner's paradox (Fechner, 1860b; Levelt, 1965b; Lehky, 1983). Fechner's paradox occurs when unequal luminances are presented to the two eyes. Under some circumstances, closing the eye with the dimmer luminance can produce the impression of increasing brightness, despite an overall decrease in the total light entering the eyes. This striking effect demonstrates that binocular combination is not a linear system[3], but requires more

[2]Similar matching experiments were also carried out by Birch (1979).

[3]A linear system has two defining characteristics; scaling and superposition. Scaling is satisfied if a given increase in input (i.e. by a factor of two) produces a proportional increase in output. Consider a simple linear system, which multiplies its input by four: an input of one unit produces an output of four, and an

a complex explanation. However, linear summation is useful as a default model, against which more elaborate models of binocular interaction can be compared (i.e. Baker, Meese and Georgeson, 2007a).

1.4.4 Dichoptic masking and binocular rivalry

Another interesting manipulation is to present a test grating to one eye, and a pedestal or masking grating to the other (dichoptic presentation[4]). Blake and Levinson (1977) showed that this configuration does produce some facilitation at low pedestal contrasts when using within-channel stimuli, although of a lesser magnitude than for monocular or binocular presentation. Small amounts of dichoptic facilitation are also apparent in the data of Levi, Harwerth and Smith III (1980) and Legge (1979). In an elegant early study, Bacon (1976) reported changes in sensitivity (*d'*) to a test when in the presence of a dichoptically presented reference (pedestal) grating. Sensitivity was greatest when test and pedestal had the same spatial properties, and reduced with changes in phase or spatial frequency, indicating that facilitation is narrowly tuned.

However, dichoptic facilitation has frequently been overshadowed by the very strong masking observed with dichoptic presentation at higher mask contrasts. Whereas the handles of monocular and binocular dipper functions have a slope of around 0.5-0.7 on logarithmic axes (Legge and Foley, 1980; Meese *et al.*, 2006), dichoptic masking produces a steeper slope, much closer to 1 (Weber's law; Legge, 1979). The masking is so strong that the test contrast must often equal or exceed the pedestal contrast before it becomes detectable.

As with the monocular arrangement, dichoptic masking becomes weaker as the pedestal differs from the test in its spatial properties (Legge, 1979). However, substantial dichoptic masking can still be obtained using orthogonal gratings (Meese and Hess, 2004). Although cross-channel dichoptic masking is a well established phenomenon, relatively little psychophysical work has been done to probe the mechanisms which cause it. This is particularly surprising, given that it is possible that these same mechanisms may play an important part in binocular rivalry (Sengpiel, Blakemore and Harrad, 1995; Levelt, 1965b), which has been the subject of numerous studies.

input of two units produces an output of eight. The ratio of the inputs (2/1 = 2) equals the ratio of the outputs (8/4 = 2). For the second characteristic, superposition, the sum of two inputs should produce the same output as the sum of the individual outputs. To illustrate using the same numbers as before, the sum of the two input values (1+2) is three, and the sum of the two output values (4+8) is twelve. Inputting the sum of the individual inputs (3), also gives an output of twelve. For an extended discussion on linearity, and its importance in systems analysis, see DeValois and DeValois (1988).

[4]Throughout this thesis, the terms monocular and binocular are used to indicate presentation of stimuli to either the one or both eyes, and dichoptic is used to describe presentation to opposite eyes. The suffix 'optic' is Greek, whereas 'ocular' is Latin. Although 'monoptic' and 'binoptic' are used occasionally in the literature, they are less familiar to most readers. There is no Latin equivalent for dichoptic, so a mixture of root suffixes is used.

Rivalry occurs when incompatible stimuli are presented to the two eyes, and cannot be fused into a single percept (see Alais and Blake, 2005, for a collection of recent work on binocular rivalry). Orthogonal gratings shown to opposite eyes are, therefore, an effective stimulus for inducing it. When viewed over a period of several minutes, observers experience alternations between one or other grating, or occasionally see a patchwork of the two images. A plaid composed of both gratings is not seen, except at very brief durations (Wolfe, 1983) or low contrasts (Liu, Tyler and Schor, 1992).

There are a number of 'top-down' explanations for this phenomenon, which relate binocular rivalry to other perceptual alternations, such as the Necker cube (i.e. Crewther, Jones, Munro, Price, Pulis and Crewther, 2005). However, it is clear that at least part of the rivalry process must occur at an early stage, prior to binocular combination, because when orthogonal gratings are shown to the same eye, rivalry is not experienced. Binocular rivalry most likely exists in order to prevent the percept of an 'impossible' image, in which two objects occupy the same location in space. It seems plausible that dichoptic masking serves a similar purpose, and that the two phenomena may result from a single process (Sengpiel *et al.*, 1995; Sengpiel, Freeman, Bonhoeffer and Blakemore, 2001).

1.5 Models of binocular interaction

There have been numerous attempts to model interocular luminance and contrast vision. One early model, proposed by de Weert and Levelt (1974), suggested that binocular brightness is a weighted sum of the two monocular luminances. The weights are a ratio of the two input values, and can be further modulated by the presence of contours in one eye. This scheme, which predicts Fechner's paradox (Fechner, 1860b), is an implementation of Hering's law of constant shares, and has been revisited by some contemporary models (Ding and Sperling, 2006b; Baker *et al.*, 2007a).

For contrast perception, Legge's (1984b) quadratic summation model, described above (equation 1.2), has been very influential. Varying the exponent values can yield binocular summation ratios between 1 and 2 (Georgeson and Meese, 2005), as observed empirically (see Table 1.1). Several recent models have been proposed as extensions of Legge's work (Meese and Hess, 2004; Meese *et al.*, 2006), some of which are considered in Chapter 3.

The literature contains further examples of binocular models based on other architectures, such as multiple mechanisms (Anderson and Movshon, 1989), separate channels (Cogan, 1987), feedback (Lehky, 1983; Gregson, 1989), parallel pathways (Maehara and Goryo, 2005), gain control (Ding and Sperling, 2006a,b), and binocular rivalry (Lehky, 1988; Kalarickal and Marshall, 2000; Wilson, 2003; Stollenwerk and Bode, 2003; Freeman, 2005).

Despite the many differences between all of these models, they share one thing in common. Each model was developed specifically to explain data from one series of experiments or one psychophysical task, and does not easily generalise to other experimental paradigms or data sets. Thus, rivalry models cannot predict contrast perception, and within-channel models do not encompass cross-channel effects. What is needed is a robust and general model architecture for binocular vision, which can account for human performance and perception using many combinations of stimuli in the two eyes. Clearly this is no simple undertaking, but one of the primary aims of this thesis is to develop models of binocular vision to account for as wide a range of empirical findings as possible.

1.6 Contrast vision in amblyopia

As well as studying the visual systems of normal humans, it is often helpful to consider what changes occur when the system is damaged. Amblyopia is a neurological condition which develops in childhood when the input to one eye is disturbed in some way. Common causes are strabismus (squint) or a large anisometropy (difference in focus between the eyes). If the ocular defect is present early in life, it can prevent the visual system from developing normally. Stereoscopic vision is rarely functional, and in many subjects the amblyopic eye is ignored or 'suppressed', relative to the normal eye. Importantly, if the ocular defect is corrected after a critical period (normally up to age six), for example by surgery for strabismus or corrective lenses for anisometropia, the deficits remain. This is strong evidence that amblyopia manifests itself at some point in the visual system beyond the eye.

A typical amblyope has poor contrast sensitivity in the affected eye, relative to the normal eye (Hess and Howell, 1977; Hess, 1979; Levi, Harwerth and Smith III, 1979; Levi *et al.*, 1980). This is usually more severe at the higher spatial frequencies, above 1cpd, even after compensation for any refractive error (Bradley and Freeman, 1981), and can affect discrimination performance as well as detection thresholds (Ciuffreda and Fisher, 1987; Kiper and Kiorpes, 1994). Often, sensitivity is so reduced in the amblyopic eye that binocular thresholds are governed by the sensitivity of the fellow eye, and binocular summation is not observed (Lema and Blake, 1977; Levi *et al.*, 1980; Holopigian, Blake and Greenwald, 1986). However, inhibitory interactions remain intact (Levi *et al.*, 1979, 1980; Harrad and Hess, 1992), and may even be stronger than in normal subjects (Pardhan and Gilchrist, 1992; Hood and Morrison, 2002; Harrad and Hess, 1992).

Surprisingly, noone has yet attempted to account for the pattern of deficits in amblyopic contrast vision using computational modelling. Any proposed model would need to account

for both the absence of binocular summation at threshold, and the differing levels of suppression beween the two eyes. Meese and Hess (2004) proposed a model architecture which they suggest might account for the observed pattern of results. However, this conjecture has yet to be tested. Chapter 7 seeks to address this issue.

1.7 Overview

The layout of the remainder of the thesis is now described. Figure 1.3 summarises the relationships between chapters, and Figure 1.4 provides an overview in terms of the stimulus space explored.

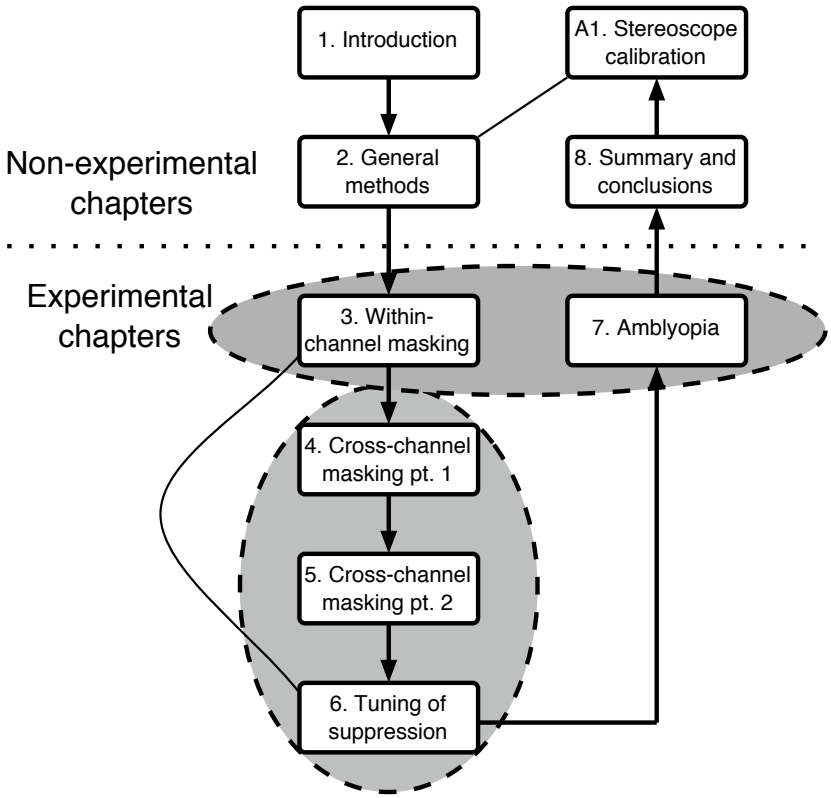

Figure 1.3: Summary of chapter layout and interrelationships. Arrows give the numerical ordering of chapters, grey elipses group chapters which comprise similar themes and experiments, and thin lines connect chapters which are otherwise associated.

Chapter 1 is this introductory chapter, which summarises relevant research, and introduces some important terms.

Chapter 2 outlines the general psychophysical approach, and gives specifics of the experimental procedures and equipment used throughout the thesis.

Chapter 3 reports the results of four experiments using within-channel stimuli (where test and pedestal have identical spatial properties). The first of these experiments recorded

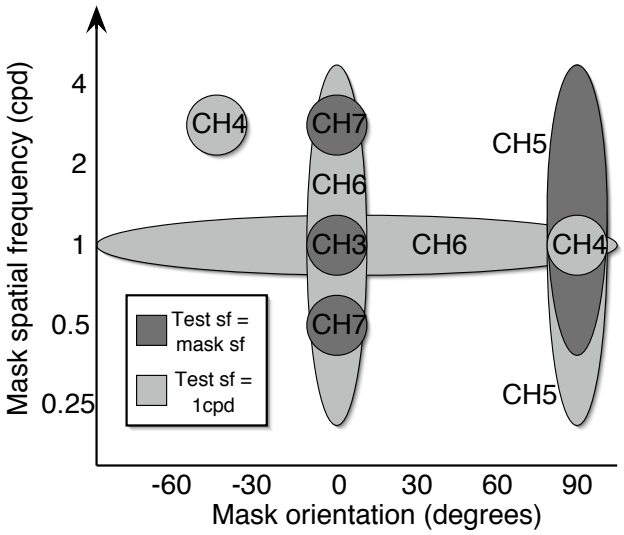

Figure 1.4: Summary of thesis as a map of stimulus space. The lozenges show the mask orientation and spatial frequency covered by various experiments. In all cases, test stimuli were horizontal (90°). Dark grey shading indicates that mask and test had the same spatial frequency. For light grey shading, the test spatial frequency was always 1cpd. In general, experiments explored both monocular and dichoptic presentation, and often several other ocular configurations. CH=chapter.

monocular, binocular and dichoptic masking for three subjects. Several models of binocular combination are considered, two of which are able to account for the empirical results. The remaining three experiments i) probe unexplored regions of the stimulus space, ii) extend the model predictions to contrast matching, and iii) assess a prediction of the models that high-contrast dichoptic masks can produce a non-monotonic psychometric function.

Chapter 4 considers cross-channel stimuli, and measures masking in several ocular configurations, at different mask contrasts and stimulus durations. Monocular and dichoptic masking are shown to arise from different mechanisms, which have different spatial and temporal properties. The two surviving models from the previous chapter are extended to account for the results. This results in the rejection of one of the model architectures.

In Chapter 5, four further cross-channel experiments are reported, inspired by recent findings from neurophysiology. The first of these tests the order in which monocular and dichoptic masking occur, and suggests that monocular masking takes place earlier in processing than dichoptic masking. In the second experiment, observers are adapted to the mask stimulus, and masking is measured. Dichoptic masking is reduced by this procedure, implying a cortical locus. The third and fourth experiments explore the spatial and temporal properties of cross-channel masking. Monocular masking is shown to be broadly tuned in both dimensions, and its magnitude is dependent on stimulus speed (TF/SF). Dichoptic masking is less broadly tuned, and is scale invariant in space and time.

Chapter 6 begins by comparing parallel dichoptic masking when mask and test are ei-

ther in-phase or in antiphase. The antiphase condition produces weaker masking and no facilitation, consistent with a version of the model in which binocular summation between channels does not occur. The antiphase paradigm also suggests an explanation for the unusual shape of dichoptic spatial frequency tuning functions reported by Legge (1979). These functions are shown to be the envelope of a broadly tuned suppressive process, and narrowly tuned binocular summation. Similar results are found for orientation tuning. The model is adapted to describe these data and summarise the tuning of the two processes.

Chapter 7 extends the contrast discrimination experiment described in Chapter 3 to strabismic amblyopes. The amblyopic data resemble those for normal subjects, but show increased thresholds in the amblyopic eye and weaker monocular facilitation. Binocular summation was found to be intact in amblyopes when a summation paradigm was used to control for sensitivity differences between the eyes. The results are considered in light of the model, and also compared to data from normal subjects with a neutral density filter attenuating the input to one eye.

Chapter 8 summarises the thesis, draws some conclusions, and suggests avenues for further study.

Appendix A outlines a subjective calibration procedure for mirror stereoscopes which was developed for some of the experiments reported here. Guidelines for accurately calibrating stereoscopes are not available in the literature.

1.8 Collaborative work and publication

The work in this thesis is my own, and is presented in this form for the first time. There is one exception to this, regarding the models presented in Chapter 3. The Two Stage model was originally formulated by Tim Meese, who also developed the Twin Summation model from that of Maehara and Goryo (2005). Details of both models were first reported by Meese *et al.* (2006). Since the majority of the thesis concerns the development of these models, it is important to include the rationale behind their inception. Furthermore, the analysis contained in the thesis is subtly different; whereas the Meese *et al.* (2006) study modelled the data averaged across subjects, here the model is fit to the data from each subject individually.

Much of the work presented in this thesis has also been published, submitted for publication, or presented at conferences and published in abstract form. During the publication process there is inevitably considerable input from co-authors. However, it should be stressed that the published reports of experiments are distinct from those contained in the thesis, for which the text, figures and computational work are entirely my own. In the following

paragraph, publications which contain data from this thesis are detailed.

The results of Experiment 1 (Chapter 3) are also presented by Meese *et al.* (2006) and form a part of two published abstracts (Georgeson, Meese and Baker, 2005; Meese, Georgeson and Baker, 2005a). Experiments 2 and 3 (Chapter 3) are published in paper (Baker *et al.*, 2007a) and abstract (Baker, Meese and Georgeson, 2005) form. Much of the analysis, and some of the figures, are duplicated here, and represent original work. The experiments reported in Chapter 4, along with Experiment 8 (Chapter 5), are also reported in Baker, Meese and Summers (2007c), and published in abstract form (Baker and Meese, 2006a,b). The thesis contains additional data from these experiments which have not yet been published. Also, the analysis presented here uses a different version of the model, and is more comprehensive than that in the paper. Experiment 10 (Chapter 5) is published in abstract form (Baker, Meese, Patel and Sarwar, 2007f), and the main experiments of Chapter 6 (but not all of the modelling) are published as both a paper (Baker and Meese, 2007) and an abstract (Baker, Meese and Patryas, 2007b). Finally, Chapter 7 has produced two papers (Baker, Meese, Mansouri and Hess, 2007d, and in preparation) and one conference abstract (Baker, Meese, Mansouri and Hess, 2007e).

Chapter 2

General methods

2.1 Using psychophysics to explore visual processing

Psychophysical methodologies have been used to study human perception for over 150 years (Fechner, 1860a). They aim to relate precise manipulations of basic stimuli to changes in a subject's performance in some task. The majority of experiments reported in this thesis use a either a detection task, in which the observer must correctly identify the presence of a signal, or a discrimination task, in which the observer must identify a change in signal magnitude. Signal detection theory (Green and Swets, 1966) provides a framework which relates the internal response generated by a signal to the probability that the observer will correctly identify its presence.

The two-alternative-forced-choice (2AFC) paradigm is favoured, as it is robust, and less prone to bias than single interval methods (though see Nachmias, 2006). In 2AFC experiments, the target signal is randomly assigned to only one of two (usually sequential) presentations, and the observer must indicate which presentation he or she believes contained the signal. Uninformed guessing, and conditions in which the signal is far below detectability, result in a chance performance level of 50% correct over multiple repetitions. As signal strength is increased, the task becomes easier, and performance tends towards 100%, producing a smooth sigmoidal curve, termed the psychometric function. A threshold is then estimated from the curve at some arbitrary level (often 75% correct), which is taken as the point at which the signal can be reliably detected.

These techniques can be repeated with various stimulus configurations, and thresholds compared across conditions. By taking many measurements (typically several hundred per condition), very precise threshold estimates can be made, and inferences can be drawn about an observer's internal state. Often, competing theories about the neural mechanisms under

study are implemented as computational models and their predictions compared (see Marr, 1983 and Graham, 1989 for the general approach). Such models are a useful tool, as they offer quantitative criteria for assessing whether a given model is a plausible approximation of what the visual system actually does.

This combination of psychophysics and computational modelling provides several distinct advantages over other methodologies currently used to study vision. It is less invasive than neurophysiological methods, such as single-cell recording, and so can be carried out using human subjects. It is also arguably a more appropriate method of studying the functional architecture of visual processing than current brain imaging techniques, such as magnetic resonance imaging (fMRI). Such techniques, which are still in their infancy, are ideal for investigating the structural architecture of vision (where things occur), but are often less informative about the computational process involved (how things occur).

2.2 Stimuli

Stimuli in all experiments were sinusoidal luminance modulations about a mean level of luminance (hereafter simply termed gratings). Their luminance profile is given by:

$$L_{x,y} = L_0(1 + C \cdot sin((cos(\theta) \cdot 2f\pi x) + (sin(\theta) \cdot 2f\pi y) + \phi)), \tag{2.1}$$

where L is luminance, L_0 is mean luminance, x and y refer to spatial locations, θ determines orientation, f determines frequency and ϕ, phase. C is Michelson contrast, as described below in equation 2.2.

Each grating was confined spatially by a circular window whose edges were tapered to zero by a half-period of a sine wave. The diameter at half-height was 4°, and the smoothing half-period was 1°. An example grating is given in Fig 2.1, along with its luminance profile. Gratings are standard stimuli for studying the visual system because their properties can be easily manipulated, and they are narrowband in the fourier domain (see Graham, 1989). Additionally, the receptive fields of many neurones in visual cortex respond optimally to sinusoidal modulations.

When studying binocular combination of signals, it is essential that stimuli presented to opposite eyes are correctly in register. If vertically oriented stimuli are used, then a small (and perhaps involuntary) horizontal vergence movement of the eyes could easily bring the two stimuli out of phase with each other. This is particularly a problem at high spatial frequencies (see Green and Blake, 1981). Therefore, except where stated, test stimuli were always horizontal gratings, at one cycle per degree of visual angle, windowed by a raised

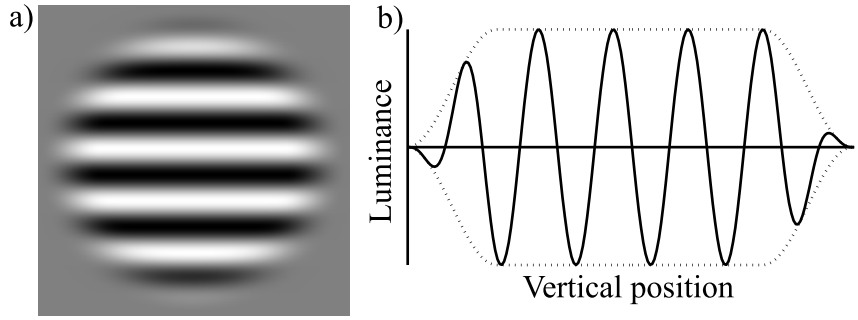

Figure 2.1: Example of a horizontal sinusoidal grating with a raised cosine spatial window. a) shows the grating, and b) gives the luminance profile of the sinusoid (dark line) and the window (dotted line). In this example, as with most stimuli used throughout the thesis, the window has a central plateau of three complete carrier cycles, and a total extent of five cycles. The full width at half-height is therefore 4 degrees.

cosine envelope with a cosine half-period of one degree.

2.3 Frame interleaving and contrast scaling

The experiments carried out for this thesis required precise manipulation of stimulus contrast. Michelson contrast is calculated as the luminance modulation around a mean value, according to the formula,

$$C = \frac{L_{MAX} - L_{MIN}}{L_{MAX} + L_{MIN}}, \tag{2.2}$$

where L_{MAX} and L_{MIN} are maximum and minimum luminances. Michelson contrast is expressed throughout as a percentage ($C\% = C \times 100$). As the contrast response of the visual system is approximately logarithmic, contrast is converted into decibels (dB) by,

$$C_{dB} = 20 log_{10} C\%, \tag{2.3}$$

where $C\%$ indicates contrast expressed as a percentage. On this scaling, 1% contrast is 0dB, and increments of 6dB correspond to a doubling of contrast (i.e. 2%=6dB, 4%=12dB). The dB metric is thus used for convenience, and for consistency with much of the existing literature. Since on a logarithmic scale absolute zero does not exist ($-\infty$), this value is labelled as 0% on figure axes which are otherwise in dB.

The equipment used was typically only able to display 256 different grey levels in a given frame. If the same grey levels are used for all stimuli, there is insufficient contrast resolution to display gratings across the visible contrast range without introducing artifacts. Look-up

tables (LUTs) were used to allocate the ideal grey levels for the desired stimulus contrast, allowing the full dynamic range of the monitor to be utilised. This is often referred to as pseudo 14-bit mode (or 12-, or 15-bit, depending on the total number of grey levels available in the graphics system), because over multiple frames all possible grey levels can be shown, even though only 256 of them are present in each frame.

LUT-based scaling presents a problem if two stimuli with widely differing contrast levels must be shown at once (as is common in masking paradigms). A frame interleaving technique was used, in which the two stimuli are presented on alternate frames, each with its own LUT. When used in conjunction with the shutter goggles (described below), this technique also allowed presentation of different stimuli to each eye. Note that frame interleaving can produce a halving of effective contrast. This was taken into account in the experimental software, so that all reported values represent the appropriate contrast level seen by the observer.

2.4 Equipment

2.4.1 Computing hardware

Several systems were used for the experiments reported. All of the specialised graphics hardware was purchased from Cambridge Research Systems, Ltd. (Kent, UK). The primary system was a ViSaGe stimulus generator driven by a Dell PC, and this was used for all experiments, except where stated. A BITS++ box, a VSG 2/4 and two VSG 2/5s were also used for certain experiments. All data analysis and computational modelling was carried out using a 1.5GHz G4 Powerbook (Apple Computer, USA). The majority of experiments took place at the vision research labs at Aston University, Birmingham, UK. The experiments in Chapter 7 took place at McGill Vision Research Unit, Montreal, Canada.

2.4.2 Monitors

All experiments used Clinton Monoray monitors, of which there were four, attached to various different systems (only one monitor was used at once). These monitors were chosen because they have a very fast phosphor decay rate ($\sim 400\mu s$), which allows them to be used in conjunction with shutter goggles (see below) while producing minimal crosstalk. Their mean luminance output is also relatively high ($\sim 200cd/m^2$), which is fortuitious, given that the goggles act as a neutral density filter with a strength of around 0.9 log units. The monitors were driven at 120Hz (except where stated) and were carefully gamma corrected using standard techniques to compensate for the inherent nonlinearity of CRT displays.

2.4.3 Shutter goggles

FE-1 shutter goggles (Cambridge Research Systems, Ltd., UK) were used to achieve presentation of different images to each eye. These contain a ferro-electric material which can be switched between opaque and transparent by reversing the polarity of an applied current. By synchronising the alternations with the CRT refresh rate, each eye can be made to view only every other frame of the display. Using the frame interleaving technique described above, this means that each eye can be shown a different image.

At the Aston lab, the goggles were mounted on a head- and chin-rest, which was fixed at the appropriate viewing distance. At the McGill lab, goggles were attached to the observer's head by means of an elasticated strap.

Goggles do not require calibration for different observers, and using them feels very natural compared to using a stereoscope because vergence is automatic. However, because they use frame interleaving, they are not ideal for experiments in which stimulus variables other than image contrast were manipulated. For such experiments, a mirror stereoscope was used.

2.4.4 Mirror stereoscope

A four-mirror stereoscope was used, and is shown as a schematic diagram in Figure 2.2. Fusion was aided by the presence of a circular stimulus aperture for each eye (9° diameter), set at the mean luminance level, within which all stimuli were displayed. Outside of the apertures, the screen was set to black. This provided a very strong cue to fusion, and was an obvious indicator if fusion was ever lost.

In addition, the device was carefully calibrated for each observer, using the procedure described in Appendix A. When calibrated, fusion was effortless, and did not require the observer to make vergence movements. This prevented any fatigue associated with keeping the eyes converged for long periods of time. The stereoscope used front silvered mirrors, and so did not suffer from internal reflections. It was constructed from optical bench components (OptoSigma, USA) according to specifications on Randolph Blake's website[1].

As mentioned above, when using goggles, the high luminance output of the monitor was attenuated by around 0.9 log units. In order to keep the luminance level the same for experiments using the stereoscope, a 0.9 log unit neutral density (ND) filter was incorporated into the stereoscope, as shown in Figure 2.2. This meant that mean luminance was constant for both stereoscope and goggle experiments, enabling a more direct comparison of results.

[1] http://www.psy.vanderbilt.edu/faculty/blake/Stereoscope/stereoscope.html

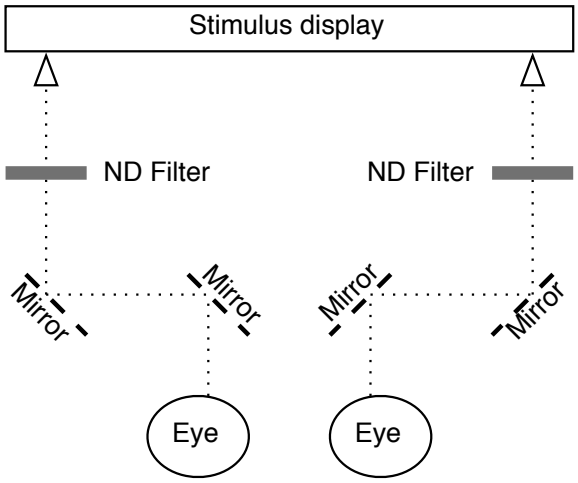

Figure 2.2: Schematic diagram of the mirror stereoscope. All four mirrors could be adjusted independently in both the horizontal and vertical planes. ND filters were used to attenuate the luminance for consistency with experiments using goggles. The dotted lines show the path the light takes between stimulus display and eye when fixation was straight ahead. In an experimental situation, the eyes converged on a central fixation point.

2.5 Experimental software

The majority of experiments reported here used the Liberator software. Developed by Rob Summers and Tim Meese, Liberator is a versatile application designed to run psychophysical experiments using the VSG series of stimulus generators, and implemented in Delphi. An extensive user interface allows the implementation of numerous experimental paradigms without the need for programming. This is, in part, responsible for the large number of experiments carried out for this thesis, as little time was spent by the author writing experimental software.

For two experiments alternative software was used because of a change in operating platform (Apple Macintosh rather than PC). For Experiments 1 and 4 (Chapter 3), Matlab code written by Mark Georgeson was used.

2.6 Data analysis and modelling

Psychometric data were analysed using a Matlab implementation of probit analysis (Finney, 1971), translated by the author from Pascal. Modelling software was also developed in Matlab. The built in *fmins* (or *fminsearch*) function was used extensively for multidimensional function minimisation. This function is an implementation of a downhill simplex algorithm (Nelder and Mead, 1965) which is commonly used to determine the best fitting model parameters for a given data set.

Goodness of fit of computational models was assessed by the widely used root mean square (RMS) error statistic. Calculated in decibels (dB; see equation 2.3), this is defined as,

$$RMSe = \sqrt{\sum_{i=1}^{n}(model_i - data_i)^2/n}, \tag{2.4}$$

where *model* and *data* are the model predictions and empirical data points, and n is the number of observations (thresholds). The smaller the RMS error, the better the fit. In general, errors <1dB are good, and errors <2dB are acceptable.

When comparing computational models, it is often useful to have some independent criterion which takes the number of free parameters into account, as these may differ between models. A commonly used statistic (i.e. Wichmann, 1999; Peirce, 2007) is Akaike's Information Criteria (or AIC; Akaike, 1974), defined as,

$$AIC = n \cdot log(RMSe) + 2p, \tag{2.5}$$

where n is the number of data points, p is the number of free parameters in the model, and *RMSe* is as defined by equation 2.4. Lower AIC scores indicate a more efficient model.

2.7 Principal Observer

The author (DHB) served as the main observer in all experiments. DHB is emmetropic, has excellent stereoacuity, and was aged between 22 and 24 during the course of all experiments. A standard sight test was carried out near the start of the PhD, which revealed a very slight astigmatism in one eye (<1 diopter cylinder). This is too small to cause substantial sensitivity differences between left and right eyes for the low spatial frequencies used here. This was further confirmed by eye-based analyses of experimental results. Vision was otherwise good, with no optical correction required.

2.8 Procedure

All experiments took place in a darkened room. Subjects used a chin and head rest, located at the appropriate viewing distance, and with any specialist equipment (such as a stereoscope or stereo goggles) mounted on the head rest (in the Montréal lab, head and chin rests were not used). Naïve subjects were given an initial briefing, and had the opportunity to complete

some practise trials before data collection began.

All experiments used a two interval, two alternative forced choice design (2IFC/2AFC). Each interval was marked by an auditory tone to reduce uncertainty about stimulus onset. For experiments where a range of mask contrasts was used, these were presented in separate blocks. Blocks occurred in a random order, determined by the computer, with subjects given the opportunity to rest between blocks. This meant that successive repetitions of an experiment would have a different block order for mask contrasts.

A staircase procedure (Levitt, 1971; Cornsweet, 1962) tracked thresholds for each condition. Three consecutive correct responses elicited a reduction in stimulus contrast, and one incorrect response caused the level to increase (3-down, 1-up rule). In most experiments, separate staircases measured performance for the left and right eyes, and were interleaved randomly. Except where stated, a staircase terminated after the lesser of 70 trials or 12 reversals. The step size was 12dB until the first reversal, and 3dB thereafter, except where stated otherwise.

Psychometric staircase data were pooled across all repetitions of an experiment, and a single threshold estimated using probit analysis (Finney, 1971). Probit analysis is a well established technique, and simulations on real and synthetic data found it to be more generally stable, and less prone to bias from outlying data points, than other methods of analysis (such as the commonly used *psignifit* software). Probit analysis fits a cumulative Gaussian function to the data, defined by the equation,

$$P(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} exp\left(-\frac{(x-\alpha)^2}{2\sigma^2}\right) dx, \tag{2.6}$$

where x is the contrast of the test stimulus, in dB units, and α and σ are free parameters which determine the midpoint (threshold at 75% correct) and spread (slope), respectively, of the cumulative Gaussian. The cumulative Gaussian is scaled appropriately for nAFC data according to,

$$\Psi(x) = \gamma + (1-\gamma)P(x); \qquad \gamma \leq \Psi(x) \leq 1, \tag{2.7}$$

where γ is given by $1/n$, and *P(x)* is given by equation 2.6. An example probit fit to psychometric data is shown in Figure 2.3. Throughout the thesis, error bars in figures typically give the standard errors of probit fits.

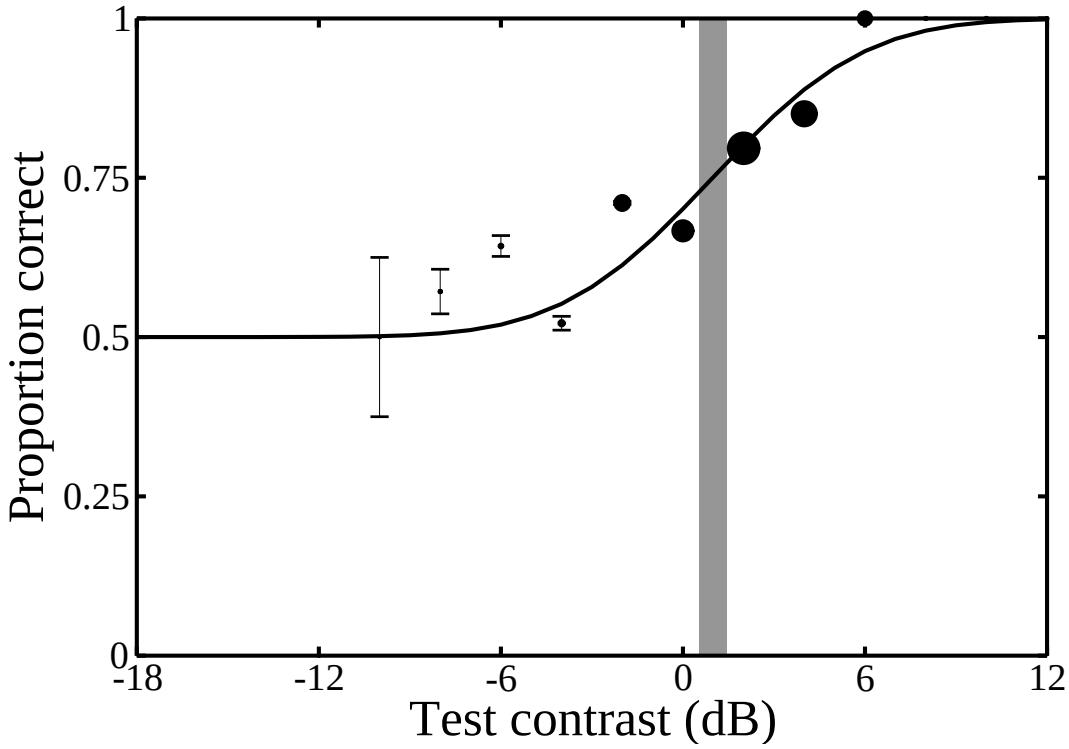

Figure 2.3: Example probit fit to psychometric data. Data are for a detection threshold from Experiment 1 (Chapter 3) for subject DHB. Symbol size is proportional to number of trials (total n=600). Error bars give binomial error, and are often smaller than symbols. The curve is the cumulative Gaussian fit calculated by probit analysis, and the shaded region defines the standard error of the threshold estimate.

Chapter 3

Within-channel masking

3.1 Experiment 1: dipper functions

3.1.1 Introduction

Although contrast discrimination functions have been reported using monocular (Legge, 1984a), binocular (Legge and Foley, 1980) and dichoptic (Legge, 1979) pedestals, no data set exists in which all three conditions were run on the same group of subjects. This makes accurate quantitative comparison of different models problematic, and suggests a worthwhile starting experiment.

During the course of this work, Maehara and Goryo (2005) published the results of a similar experiment, which examined these three conditions. However, their masks were spatially more extensive than their test stimuli, so the results are not directly comparable to previous work (Legge, 1979, 1984a; Legge and Foley, 1980) or the paradigm used here. The model which they present to account for their data is discussed in Section 3.1.4.

3.1.2 Methods

Apparatus and Stimuli

Stimuli were 1cpd horizontal gratings, constructed as described in Chapter 2. Ferro-electric goggles allowed presentation of different contrasts to the two eyes, using a frame interleaving technique. Stimulus presentation was governed by a Bits++ box (Cambridge Research Systems, Ltd., UK) running in mono mode, and connected to an Apple Macintosh G4, running Matlab 5.2, and using the PsychToolBox software (Brainard, 1997; Pelli, 1997). All other variables were as described in Chapter 2.

Procedure

Subjects were seated in a darkened room, with their head in a support, to which the goggles were attached. The viewing distance was 114cm. A 3-down 1-up staircase routine (Cornsweet, 1962) was used to govern test contrast level. On each trial of a 2AFC procedure, observers viewed two intervals, one containing the pedestal with a test increment, and the other containing the pedestal alone. The task was to indicate with a button press, which of the two intervals contained the test increment (stimulus duration = 200ms; ISI = 500ms).

Stimuli were presented in three ocular configurations of pedestal and test. In the monocular condition, both were presented to the same eye, in the dichoptic condition they were presented to opposite eyes, and in the binocular condition both were presented to both eyes. Two independent staircases were run for each condition, to allow counterbalancing across eyes. All subjects completed 6 replications of the experiment, each lasting around 90 minutes. Stimuli were blocked by pedestal contrast and, within each block, all ocular conditions and eye-of-presentation staircases were interleaved. This meant that subjects were not aware of which condition, or eye, was being tested on any given trial. The data were then collapsed across eye of presentation and replication (around 600 trials per threshold), and thresholds at the 75% correct point were estimated using probit analysis (Finney, 1971).

Observers

Three experienced psychophysical observers, male and aged between 22 and 30, took part in the experiment. All wore their normal optical correction, and had no abnormalities of binocular or stereo vision.

3.1.3 Results

Several interesting features are apparent in Figure 3.1, which shows the results for all three subjects. The characteristic dipper shape is observed in all three conditions, consistent with other studies on contrast discrimination (Nachmias and Sansbury, 1974; Legge and Foley, 1980; Blake and Levinson, 1977; Wilson, 1980). The dichoptic condition shows weaker facilitation than the other two conditions (around 4dB), and much stronger masking at higher pedestal levels (Legge, 1979). Binocular summation ratios at threshold were greater than $\sqrt{2}$ for all three subjects (DHB=1.54, DJH=1.52, RJS=1.83). The monocular and binocular dipper handles converge at higher pedestal levels (consistent with Legge, 1984a), removing the binocular advantage for contrast discrimination.

The presence of facilitation in the dichoptic condition is of particular theoretical interest,

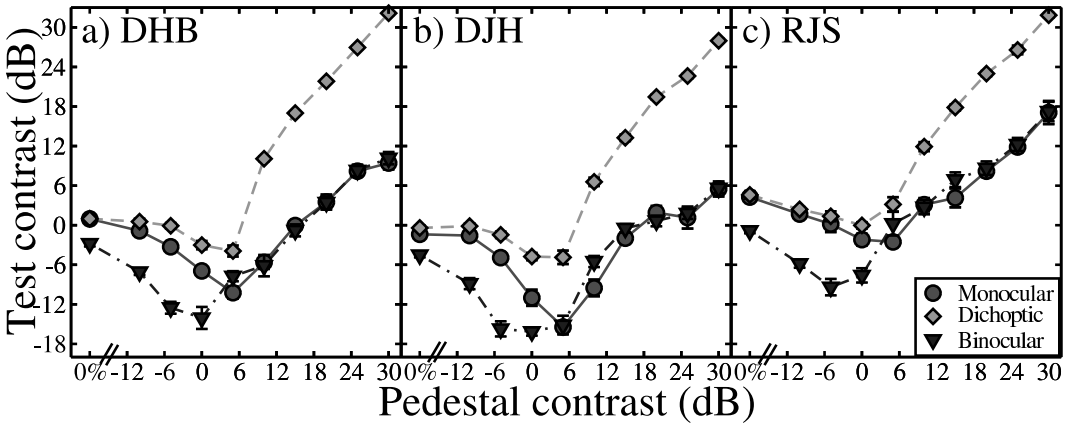

Figure 3.1: Results of Experiment 1 for three subjects, showing contrast discrimination functions for monocular, dichoptic and binocular presentation of a horizontal 1cpd grating, as described in the text. Data were pooled across eye, and over six repetitions of the experiment, such that each data point is estimated from around 600 trials. Error bars give the standard error of the probit fit, and in most cases are smaller than symbols. The leftmost points in each panel correspond to detection thresholds (0% pedestal contrast).

as it has rarely been reported in the literature (though see Blake and Levinson, 1977; Levi *et al.*, 1980) and is not predicted by many models of binocular contrast processing. Assuming an accelerating transducer function, the facilitation could be caused by within-channel stimulation from the pedestal. Alternatively, it could be a result of uncertainty reduction or some other process (Pelli, 1985; Petrov, Verghese and McKee, 2006). Bird *et al.* (2002) point out that within-channel facilitation produces a characteristic reduction in the slope of the psychometric function. If this occurred in the dichoptic condition here, it would suggest that within-channel stimulation was responsible for dichoptic facilitation, rather than any other explanation. This is investigated in the following section.

Psychometric slope data

Quick (1974) first introduced the Weibull function as an informative model of the psychometric function. The Weibull function is mathematically related to the probability summation model of vision (Sachs, Nachmias and Robson, 1971), in which an observer monitors a number of stochastically independent noisy channels, and bases their decision on whether or not at least one channel is above threshold. According to this model, as a stimulus activates more detectors (e.g. by increasing in area, or containing multiple components) the probability of one detector exceeding its threshold will increase, and performance will improve. Although this model relies on High Threshold Theory, which has been discredited (Nachmias, 1981; Tyler and Chen, 2000), Weibull β values still provide a convenient measure of steepness of the psychometric function, which can be compared to estimates from previous studies.

The Weibull function is defined by the expression,

$$W = 1 - e^{-(\frac{s}{\alpha})^{\beta}}, \tag{3.1}$$

where s is stimulus strength in linear units (i.e. % contrast), α determines the horizontal placement, and β the slope of the function. Although very different from the equation for the cumulative gaussian used by probit analysis (equation 2.6), both produce sigmoidal functions which provide good fits to psychometric data, and are very similar in form. On log axes, a change in sensitivity (α) produces a lateral shift of the psychometric function for both equations, so the steepness is determined entirely by the slope parameter (σ or β). The similarity in form allows for conversion between measures of slope for the two functions; the standard deviation (σ, in dB units of contrast) of the cumulative gaussian is inversely proportional to the Weibull slope parameter (β), such that $\beta = \frac{10.3}{\sigma}$ (Baker *et al.*, 2007d).

The median psychometric slopes for all subjects are plotted in Figure 3.2. Slope estimates were calculated by performing independent probit fits to the data from each session, and taking the median slope value (n=6). While different from the method used to estimate thresholds, this procedure reduced any artifactual shallowing of the psychometric function caused by a non-stationary observer (it is well established that psychophysical thresholds vary slightly over time, i.e. Wertheimer, 1953; Hallett, 1969b,a; Home, 1978).

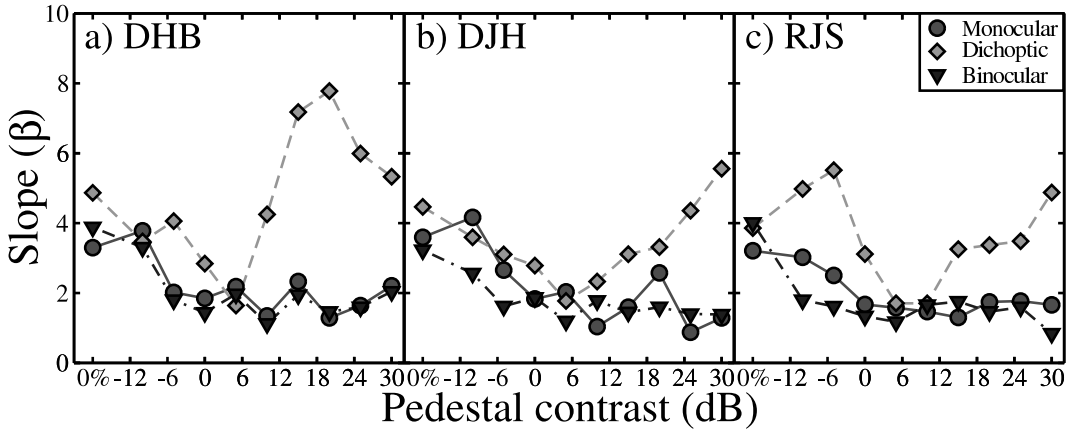

Figure 3.2: Psychometric slope data for Experiment 1. Panel layout is as for Figure 3.1, and slope data are median values, calculated as described in the text.

At detection threshold, slopes are steep in all conditions ($\beta \simeq 4$), consistent with previous studies (Robson and Graham, 1981; Bird *et al.*, 2002). In the probability summation model, this can be attributed to monitoring a large number of noisy detectors (Pelli, 1985). However, it is also explained by the Legge and Foley (1980) transducer model, as at low contrasts around detection threshold, the saturation constant (Z) dominates the denominator of the equation (see equation 1.1). This means that the effective exponent of the transducer is large ($p \simeq 2.4$). From signal detection theory (Green and Swets, 1966), the log-log slope

of the *d'* power function depends on the contrast transducer. For a linear transducer, *d'* increases linearly with contrast, producing a slope of 1 on log-log axes. The slope for nonlinear transducers is equal to the power of the transducer. The Weibull β parameter is proportional to the slope of the *d'* power function, such that $\beta \simeq 1.3d'_{slope}$ (Pelli, 1987; Tyler and Chen, 2000; Strasburger, 2001). An empirical $\beta = 4$ therefore implies an effective exponent of around 3 at detection threshold (although given that staircases can produce systematically steeper slopes, (see Kaernbach, 2001), the true value may be a little lower).

As expected based on previous work (Foley and Legge, 1981; Legge *et al.*, 1987; Bird *et al.*, 2002) the regions of facilitation in the monocular and binocular conditions correspond to a reduction in the slope of the psychometric function. Slopes remain shallow ($\beta \simeq 1.5$) in these conditions at higher pedestal contrasts. This could be due to uncertainty reduction (Pelli, 1985), as monitoring fewer irrelevant mechanisms will result in a shallower psychometric function. Alternatively, in the transducer model, the pedestal begins to dominate the denominator of the equation, so the effective exponent of the transducer becomes smaller ($p - q \simeq 0.4$), reducing the psychometric slope.

In the dichoptic condition, there is also a reduction in slope at around 0dB pedestal contrast, indicating within-channel facilitation. However, the slopes then become much steeper at high dichoptic pedestal contrasts ($\beta \simeq 6$), often greatly exceeding the levels observed at threshold (where $\beta \simeq 4$). This surprising pattern of slope data is not predicted by either the basic transducer model or the uncertainty model, and will be a challenge for candidate models of binocular contrast discrimination to replicate.

3.1.4 Models

Four models are assessed in this section, and are outlined below, with model diagrams shown in Figure 3.3. For all models, best fitting parameters were derived from 100 runs of a downhill simplex algorithm (Nelder and Mead, 1965), set to minimise the root mean square (RMS) error between model and data. Threshold predictions were generated according to the formula,

$$k = resp_{(C+\Delta C)} - resp_{(C)}, \tag{3.2}$$

where *resp* denotes model response, C is pedestal contrast, ΔC is test contrast, and k is a free parameter in the model. The parameter k also represents the standard deviation of the intrinsic (fixed) noise of the system.

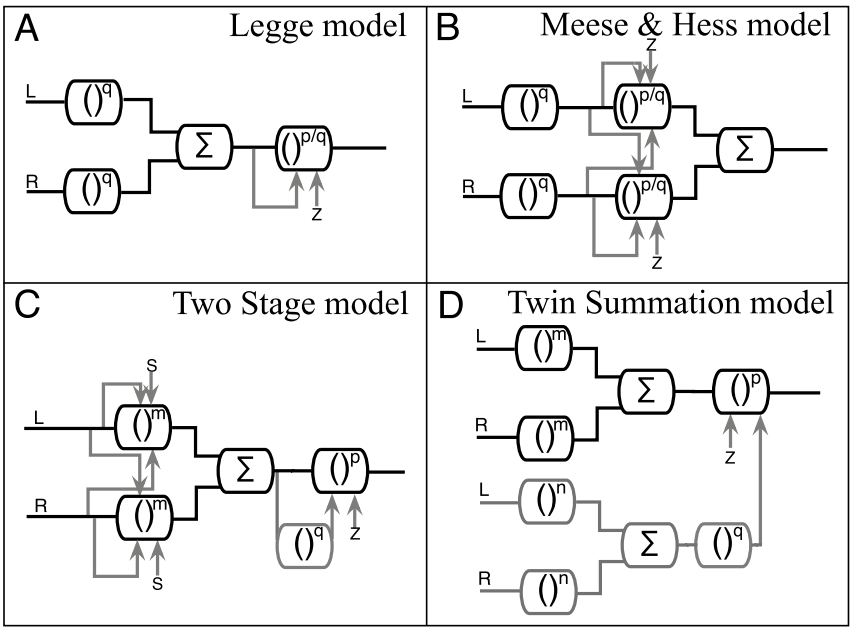

Figure 3.3: Schematic diagrams of four model architectures. In all cases, L and R correspond to left and right eyes. Black pathways are excitatory, and grey pathways inhibitory. Arrows indicate division, and brackets raised to a power denote exponentiation of inputs. Σ represents summation of inputs. Exponents (*m, n, p, q*) and saturation constants (*S, Z*) are free parameters in the models.

Legge model

The first model uses an architecture similar to that described by Legge (1984b). Contrast signals are raised to an exponent, q, before binocular summation, and are then subject to a gain control type nonlinearity (Meese and Hess, 2004; Wilson, 1980), which extends Legge's formulation to suprathreshold conditions. Figure 3.3A shows a schematic diagram. The full model equation is,

$$resp = \frac{({C_L}^q + {C_R}^q)^{p/q}}{Z + {C_L}^q + {C_R}^q}, \tag{3.3}$$

where C_L and C_R are input contrasts to the left and right eyes respectively, and p, q and Z are free parameters in the model. Figure 3.4A-C shows the model fits, with parameters and error statistics given in Table 3.1.

Whilst the model provides a reasonable fit to the data with only four free parameters, it is lacking in several key features. It cannot account for the observed dichoptic facilitation, and produces insufficient binocular summation at threshold. It also slightly underestimates the magnitude of dichoptic masking at high pedestal contrasts for all three observers. As Meese *et al.* (2006) have shown, it is possible to constrain the model parameters to provide the correct amount of binocular summation, but at the expense of disrupting the rest of the functions. In sum, this version of Legge's model is unable to account for the present data.

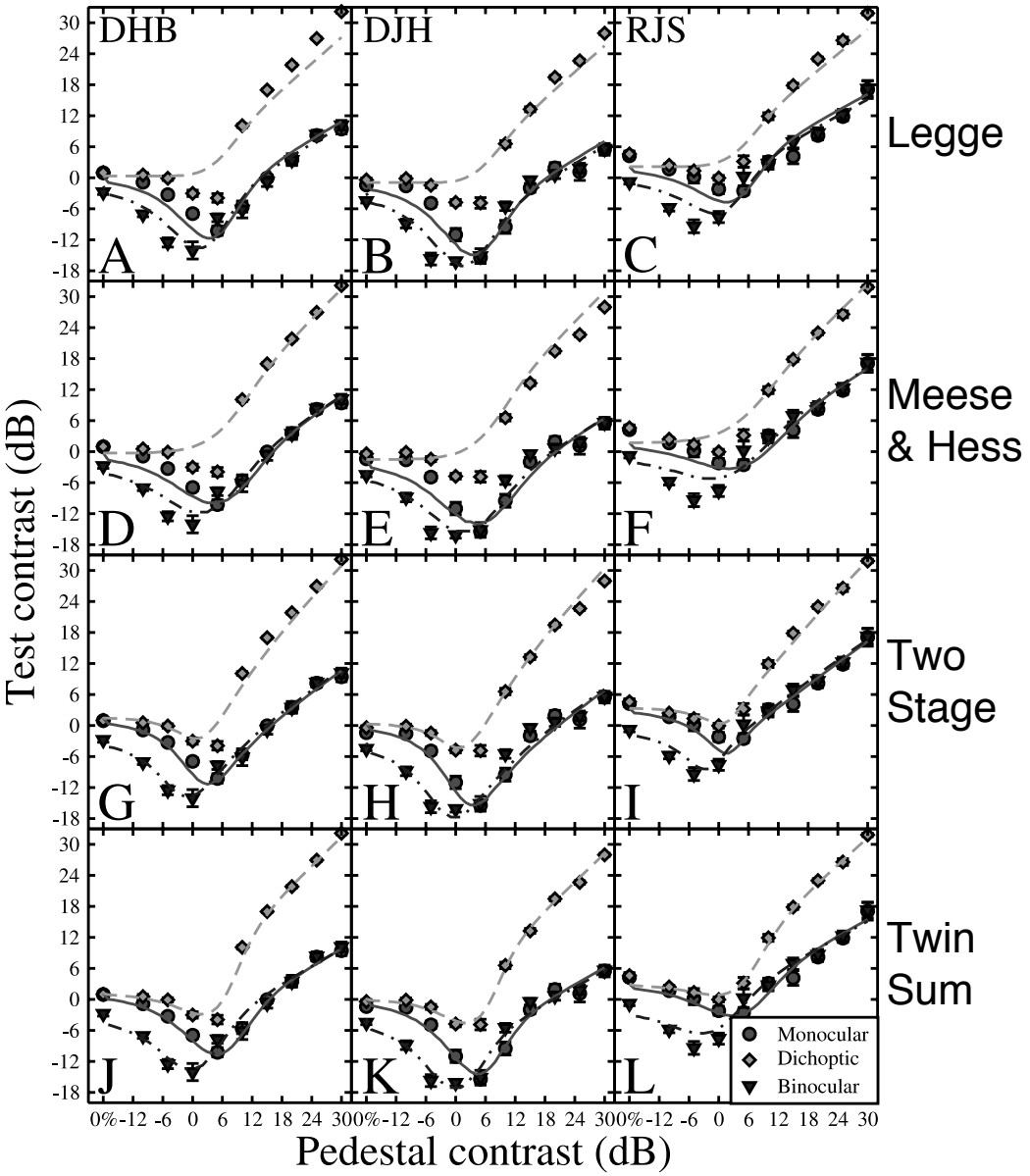

Figure 3.4: Best fits of four models to the data of Experiment 1. Data are replotted from Figure 3.1, and curves show the model fits. Panels A-C show the Legge model, D-F show the Meese & Hess model, G-I show the Two Stage model, and J-L show the Twin Summation model. Parameters and error statistics for all models are given in Table 3.1.

Model	Subject	Figure	RMSe (dB)	AIC	Slope RMS (β)	m	n	S	p	q	Z	k
Legge	DHB	3.4A	2.49	35.37	1.18	-	-	-	3.44	3.02	5.06	0.19
Legge	DJH	3.4B	2.18	31.38	0.64	-	-	-	3.33	2.92	4.93	0.13
Legge	RJS	3.4C	2.04	29.39	0.87	-	-	-	3.20	2.76	3.36	0.42
Meese & Hess	DHB	3.4D	2.03	29.24	1.11	-	-	-	2.54	2.16	5.27	0.15
Meese & Hess	DJH	3.4E	2.56	36.20	1.94	-	-	-	2.74	2.33	5.03	0.11
Meese & Hess	RJS	3.4F	1.91	27.41	1.05	-	-	-	2.25	1.86	3.65	0.31
Two Stage	DHB	3.4G	1.39	21.88	1.09	1.19	-	0.65	8.61	6.74	0.17	0.14
Two Stage	DJH	3.4H	1.73	28.44	1.52	1.28	-	0.89	6.45	5.15	0.19	0.09
Two Stage	RJS	3.4I	1.14	15.93	0.83	1.25	-	1.28	7.84	6.35	0.02	0.28
Twin Summation	DHB	3.4J	1.11	15.13	1.10	1.40	1.21	-	2.28	2.27	7.12	0.18
Twin Summation	DJH	3.4K	1.13	15.67	1.01	1.54	1.40	-	2.28	2.18	6.46	0.14
Twin Summation	RJS	3.4L	1.21	17.72	1.03	1.22	1.05	-	2.28	2.22	4.29	0.40

Table 3.1: Best fitting parameters and error statistics for the data from Experiment 1 for four models. All values are rounded to 2 decimal places. Note that parameters have different meanings depending on the model, as described in the text. Model fits are shown in Figure 3.4.

Meese and Hess model

This model, also called the late summation model, was proposed by Meese and Hess (2004) and is shown schematically in Figure 3.3B. Here, dichoptic masking is caused by divisive interocular suppression between left and right channels, as follows:

$$resp = \frac{{C_L}^p}{Z + {C_L}^q + {C_R}^q} + \frac{{C_R}^p}{Z + {C_R}^q + {C_L}^q}, \tag{3.4}$$

where symbol meanings are as for equation 3.3. Best fits of the model are shown in Figure 3.4D-F with parameters and error statistics given in Table 3.1.

Despite the radically different architecture, this model produces very similar results to the Legge model, and has the same number of free parameters (4). The original version of this model also incorporated a dichoptic weight parameter, w. This extra degree of freedom did not greatly improve the model's performance, and adequate fits were obtained with it set to unity. Although this model fails in much the same manner as the Legge model, the architecture is more flexible, and can be adapted to accommodate the desired features. Such manipulations led to the formulation of the Two Stage model.

Two Stage Contrast Gain Control model

The failure of the Meese and Hess model to predict sufficient binocular summation stems from the strong exponentiation prior to binocular summation. At detection threshold, the denominator of the gain control equation is dominated by the saturation constant Z, because the other terms are negligible in magnitude. This means that the model response around detection threshold is largely determined by the exponent p. In the absence of gain control effects, the amount of binocular summation is directly related to this exponent, such that,

$$B_{sum} = 2^{1/p}. \tag{3.5}$$

From this equation, when p=1, summation is linear (ratio of 2), and when p=2, summation is quadratic (ratio of $\sqrt{2}$). As the empirical summation ratios all fall between these two extremes, ideal exponents for the first stage should fall between 1 and 2. However, a shallow exponent ($p < 2$) causes the model to perform badly at suprathreshold levels. Hence, the best fitting values of p derived by the simplex algorithm are all greater than 2, and provide too little summation (see Table 3.1). What is required is a model which features weak exponentiation before binocular summation, and a larger compensating nonlinearity after summation.

The Two Stage model (Figure 3.3C) fulfills these criteria. It features interocular suppression at an early gain control stage similar to that of the Meese and Hess model,

$$stage1 = \frac{{C_L}^m}{S + C_L + C_R} + \frac{{C_R}^m}{S + C_R + C_L}, \tag{3.6}$$

in which m and S are free parameters, and other terms are as described above. Typical values of m are around 1.25, which gives the required amount of binocular summation. After binocular summation, there is a second gain control stage, with a further nonlinearity,

$$resp = \frac{stage1^p}{Z + stage1^q}, \tag{3.7}$$

where p, q and Z are further free parameters. As shown in Figure 3.4G-I, this model performs well for all three subjects. The amount of binocular summation is closer to the empirical values, and the model also predicts the observed dichoptic facilitation. RMS and AIC scores are lower than for the previous two models, the latter despite an increase in the number of free parameters (to six).

To illuminate which feature provides the facilitation, consider Equation 3.7 in isolation. The input to this stage is the summed binocular signal, and as such it is agnostic about eye of presentation. It is identical in form to the contrast nonlinearity proposed by Legge and Foley (1980) to account for facilitation in binocular contrast discrimination functions and described in Chapter 1. Thus, as the dichoptic pedestal contrast increases, the input to Equation 3.7 increases, and produces facilitation at low pedestal contrasts.

The Twin Summation Gain Control model

Maehara and Goryo (2005) proposed a model to account for their results for a similar contrast discrimination experiment to that reported here, but in which the pedestal was spatially more extensive than the test. Based on the work of Foley (1994), it had the form,

$$resp = \frac{[(S_{Em}.C_{mL} + S_{Et}.C_{tL})^p + (S_{Em}.C_{mR} + S_{Et}.C_{tR})^p]^p}{Z + [(S_{Im}.C_{mL} + S_{It}.C_{tL})^q + (S_{Im}.C_{mR} + S_{It}.C_{tR})^q]^q}, \tag{3.8}$$

where S is a parameter representing sensitivity to either mask (m) or test (t) and in either the excitatory (E) or inhibitory (I) section of the equation. As above, C denotes contrast, L and R are left and right eyes, and p, q and Z are free parameters. Maehara and Goryo (2005) allowed three of the four sensitivity parameters to vary, resulting in a model with seven free parameters (their decision variable, D, which is exactly equivalent to k in Equation 3.2, was fixed at unity).

For their stimuli, the model performed well. However, when mask and test are spatially identical, as in the present paradigm, the sensitivity parameters must become equal, and the model can be reduced to,

$$resp = \frac{({C_L}^p + {C_R}^p)^p}{Z + ({C_L}^q + {C_R}^q)^q}, \tag{3.9}$$

where symbols have the same meanings as previously. As Meese *et al.* (2006) have shown, this four-parameter model (*p, q, Z* and *k*) performs at least as well as the two four-parameter models described above. However, by allowing an extra two free parameters, the model becomes better able to accommodate the data of Experiment 1. The model shown in Figure 3.3D is called the Twin Summation Gain Control Model, and is described by the equation,

$$resp = \frac{({C_L}^m + {C_R}^m)^p}{Z + ({C_L}^n + {C_R}^n)^q}, \tag{3.10}$$

where C_L and C_R represent contrasts to the left and right eyes respectively, and *m, n, p, q* and *Z* are free parameters. This version of the model also uses Equation 3.2 to determine the threshold criteria, so *k* is the sixth and final free parameter in the model.

Although very different in form from the Two Stage model, the Twin Summation model also captures all the important features of the data set (Figure 3.4J-L). It produces slightly superior fits to the data, in terms of RMS error, than the Two Stage model, and has the same number of free parameters (making a comparison based on the AIC score equivalent to one based on RMS error).

3.1.5 Predicting psychometric slopes

The four models outlined above can make predictions about the slope of the psychometric function in different experimental conditions. Signal detection theory (Green and Swets, 1966) states that signal discriminability, d' (d-prime) can be calculated by,

$$d' = \frac{R_{s+n} - R_n}{\sigma}, \tag{3.11}$$

and that d' is related to proportion correct in a 2AFC task by,

$$P = \Phi(\frac{d'}{\sqrt{2}}), \tag{3.12}$$

where *R* is the response to either *n* (noise) or *s + n* (signal plus noise), and *P* is the

proportion of correct responses. The parameter σ is the standard deviation of the noise, which is assumed to be independent of signal strength, and Φ denotes the normal integral. In the context of the models above, R_n is the model response to the pedestal alone, and R_{s+n} is the response to the pedestal and test together. The parameter k is the product of σ and the value of d' at a particular point on the psychometric function (the threshold), such that,

$$\sigma = \frac{k}{d'}. \tag{3.13}$$

Probit analysis estimates a threshold at the 75% correct point, which corresponds to a d' score of 0.95. This means that the value of σ can be estimated using the model parameter k and predict the proportion of correct responses for any combination of test and pedestal using Equations 3.11 and 3.12.

Model psychometric functions were generated for each set of model parameters in Table 3.1. The slope value was determined by the same procedure as for the empirical data, using the actual number of trials carried out at each test contrast level in the experiment. Parameter-free predictions for all subjects are shown in Figure 3.5, with the RMS error between slope data and model given in Table 3.1.

All four models predict performance well in the monocular and binocular conditions. The Legge model gives an excellent numerical fit for DJH and RJS (see Table 3.1), however it does not predict the steep dichoptic slopes at high pedestal contrasts. The Meese & Hess model does predict the steep slopes, but has a larger RMS error for all subjects than either the Two Stage or Twin Summation models. The Two Stage and Twin Summation models capture the form of the slope data well, and again emerge as the preferred models.

3.1.6 Surviving models

Only the Two Stage and Twin Summation models were able to account for the observed details in the pattern of both discrimination thresholds and psychometric slopes. They also produce a more efficient fit than the other models, using the AIC statistic. The Legge model and the Meese and Hess model can therefore be rejected. The remainder of this chapter further explores the predictions made by the two surviving models.

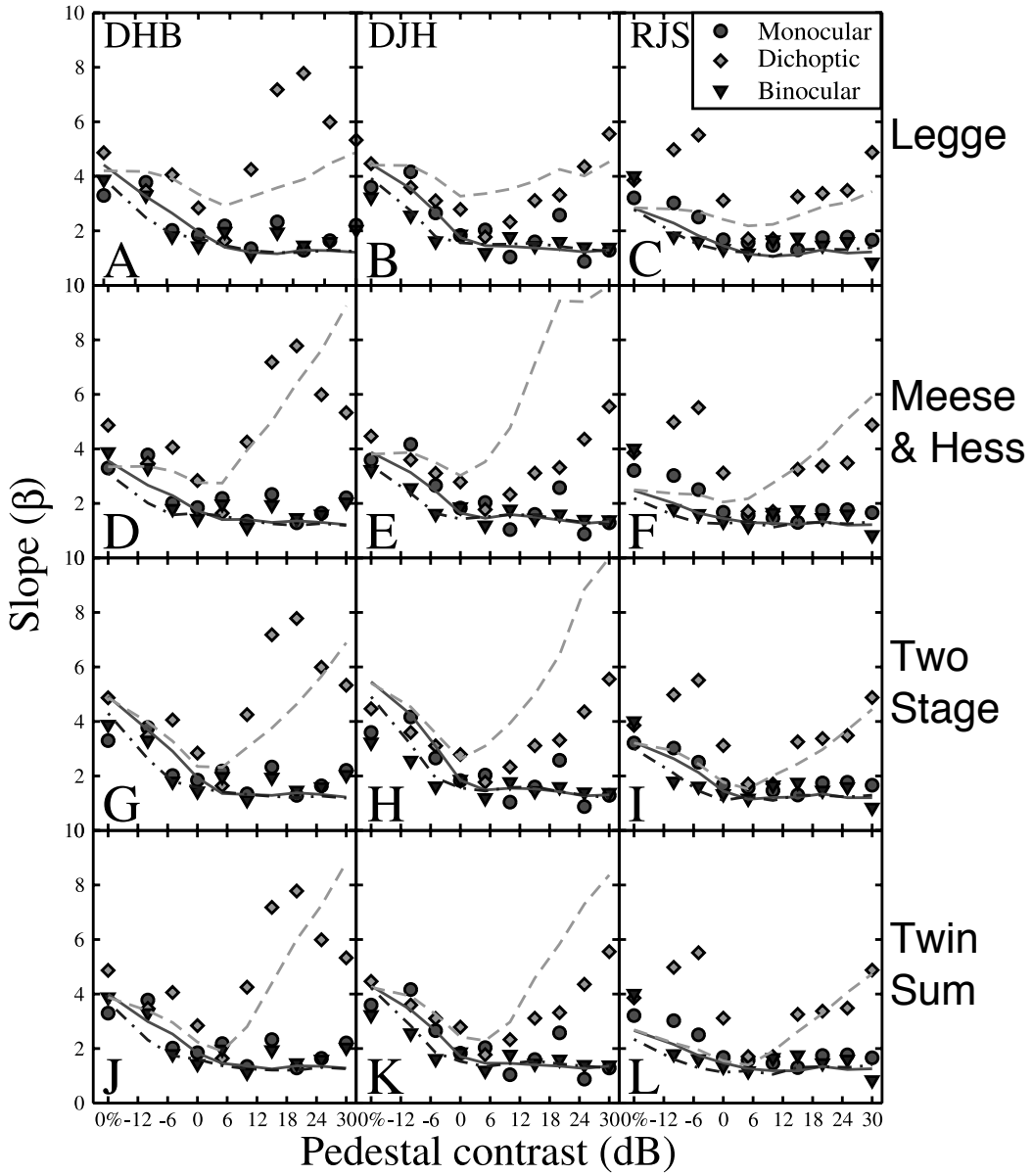

Figure 3.5: Psychometric slope predictions for Experiment 1. Data are replotted from Figure 3.2, and curves are predictions for each of the four models under discussion, based on the parameters derived from fitting the threshold data (see Table 3.1). Panel layout is the same as for Figure 3.4. Slopes are presented as Weibull β values, as described in the text.

3.2 Experiment 2: twin mask contrast discrimination

3.2.1 Introduction

Figure 3.6 shows the output of the Two Stage model for all possible combinations of input to the two eyes between -10 and 30dB, scaled in either log (3.6A) or linear (3.6B) units. The response is scaled in units of σ, calculated using Equation 3.11. The points and lines superimposed on the surface represent the conditions tested in Experiment 1. From left to right, these are dichoptic, binocular and monocular. Each point represents a pedestal level, with lines indicating the direction in which the test increments progress for the appropriate ocular configuration. Adding test increments increases the output until a threshold is reached (an increase of 0.95 σ units). Thus, discrimination performance depends on the gradient of the surface (on linear scaling) at a given point, with steeper gradients requiring smaller test increments, and thus yielding lower thresholds, than regions of shallow gradient.

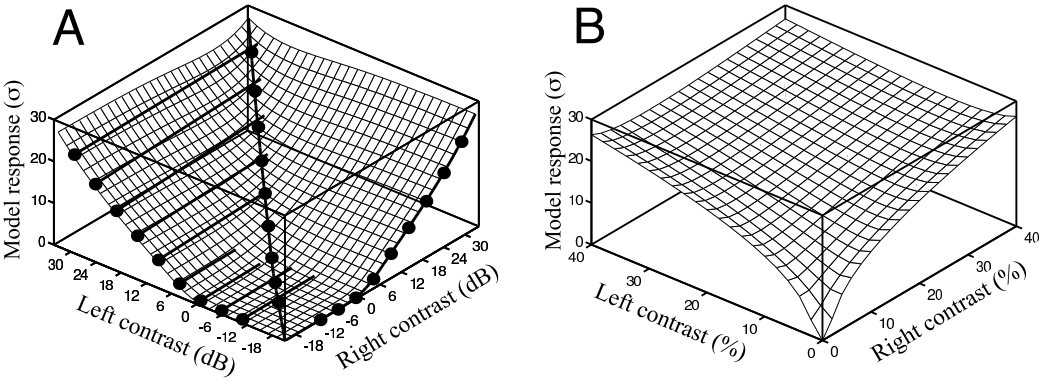

Figure 3.6: Model response plotted on logarithmic (left) and linear (right) axes. Output is scaled to units of the noise term, σ. Lines and symbols in A represent the three conditions of Experiment 1.

It is clear, however, that the conditions tested in Experiment 1 left large regions of the possible stimulus space unexplored. Experiment 2 explores how well the models can generalise across the rest of the stimulus space. This was achieved using a 'twin mask' paradigm (Foley, 1994) in which pedestals of differing contrasts are presented to the two eyes. These results have previously been published, both in paper (Baker *et al.*, 2007a) and abstract (Baker *et al.*, 2005) form.

3.2.2 Methods

Apparatus and Stimuli

Stimuli were generated using a ViSaGe framestore (Cambridge Research Systems, Ltd., UK) connected to a PC running the Liberator software. Other variables were as for Experiment 1.

Procedure

Twenty-five pairs of pedestal contrasts were generated by factorial combination of contrasts from -10 to 30dB in steps of 10dB. The test increment was presented to one eye only. For clarity, the fixed contrast grating presented to the same eye as the test increment is hereafter referred to as the pedestal, whilst the grating in the other eye will be termed the mask.

Different mask-pedestal combinations were blocked and tested in a random order. Within each block, two independent staircases were run to counterbalance across eye of presentation. The experiment was repeated six times by each observer, and the data analysed by the same method as for Experiment 1.

Subject DHB also performed an extended version of the experiment, in which the base contrasts ranged from -10 to 30dB, but in steps of 5dB, and also included a 0% level. This version of the experiment had 100 conditions, and so took considerably longer to complete. The experimental variables were also subtly different, as stimuli were all of the same phase (0°) relative to the fixation point, and staircase pairs were used instead of single staircases. These slight methodological differences appear to have had little effect on the results of the experiment.

Observers

Two of the observers who participated in Experiment 1 (DHB and DJH) also completed this experiment.

3.2.3 Results

The panels in Figure 3.7 show discrimination thresholds plotted against pedestal contrast, with mask contrast given by the symbols. As such, the data can be conceived of as monocular dipper functions in the presence of dichoptic masks of different strengths. When the mask component is weak (-10dB), the data are similar in form to the monocular functions in Figure 3.1. As the mask contrast increases, thresholds first decrease slightly (compare the leftmost

points for mask contrasts of -10 and 0dB), before increasing. This is precisely what should be expected from the dichoptic condition of Experiment 1, which showed a small amount of facilitation at low dichoptic mask contrasts, and strong masking at high contrasts.

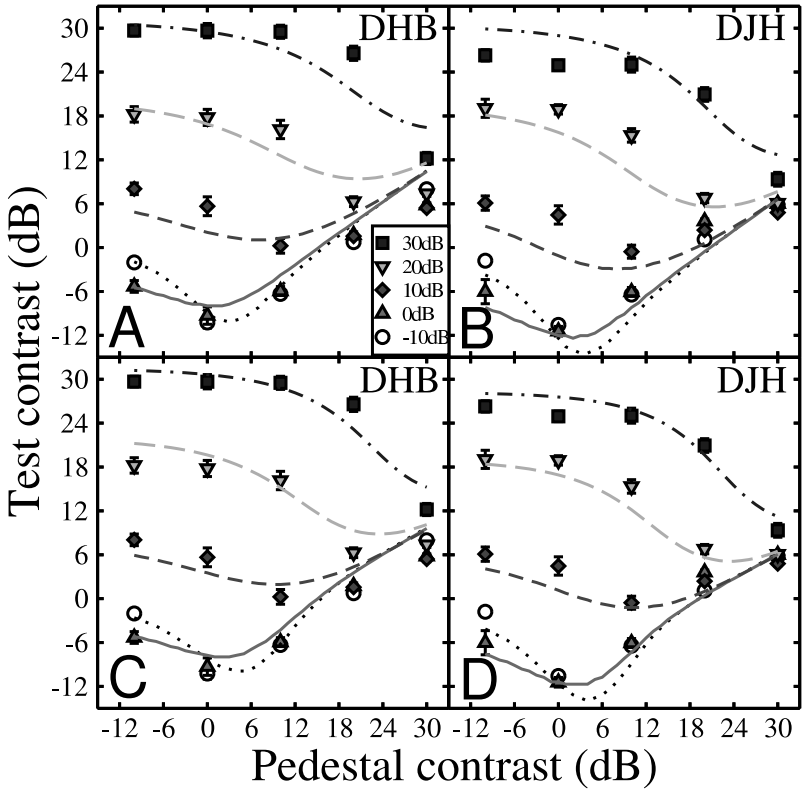

Figure 3.7: Results and model predictions for Experiment 2. The ordinate gives the discrimination threshold at different values of pedestal contrast (abscissa), and dichoptic mask contrast (symbol type). Data (symbols) are duplicated for each subject in upper and lower panels. Curves are the predictions of the Two Stage (upper) and Twin Summation (lower) models, generated using the parameters given in Table 3.1 (no free parameters). RMS errors, by panel, are A=2.82dB; B=2.75dB; C=2.28dB; D=1.77dB.

The region of 'monocular' facilitation is seen to shift upwards and to the right as mask contrast increases. This is similar to the findings of Foley (1994), who also found a diagonal shift of the dip region when an extra mask was added. However, Foley's stimuli were all binocular, and his extra mask differed in its spatial properties, so the conditions were quite different from the present experiment. Here, this shift results in an interesting effect at high mask contrasts. When mask and pedestal contrast are both 30dB, thresholds are substantially lower (by around 18dB, or a factor of eight) than when mask contrast is 30dB but pedestal contrast is -10dB. Thus, the condition with the greatest mask energy across the eyes does not confer the greatest amount of masking.

These features are also apparent in Figure 3.8, which shows the results of the extended version of the experiment plotted as a surface. The horizontal axes represent mask and pedestal contrasts, and the vertical axis gives the threshold. This representation makes the

comparison to the dichoptic condition of Experiment 1 more apparent, as a dichoptic dipper is plotted along the mask contrast axis (right to left). The translation of the monocular dip region is also apparent in the surface plot. This is most visible when the upper edges of the surface are compared; the high mask contrast edge (top left) is substantially higher than the high pedestal contrast edge (top right).

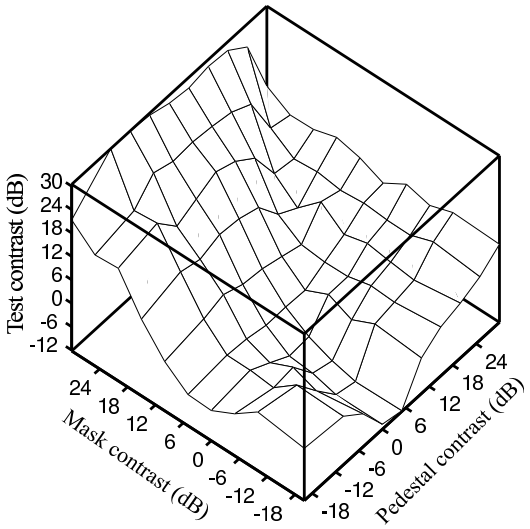

Figure 3.8: Results of an extended version of Experiment 2, in which pedestal and mask contrasts were sampled every 5dB. Data are for subject DHB, and have the same general form as the results of the main experiment. Along the 'Pedestal contrast' axis is plotted a monocular dipper function, as in Experiment 1, and the 'Mask contrast' axis shows a (reversed) dichoptic dipper. The extreme minima of each abscissa represents 0% contrast.

One discrepancy between the data in the surface, and those of Experiment 1, is the reduction in thresholds at high dichoptic mask contrasts, and 0% pedestal contrast (upper left hand corner of the plot). This may be due to the subject using a different response strategy, most likely related to the paradoxical effect discussed below in section 3.4. However, the main experiment used a different blocking procedure, which gave subjects less opportunity to change their response strategy, and apparently did not suffer from this problem.

3.2.4 Models

The data in Figure 3.7 are duplicated on upper and lower panels, with curves representing predictions of the Two Stage (upper) and Twin Summation (lower) models, using parameters derived from the fits to the Experiment 1 data (Table 3.1). The fits are reasonable, and capture all the important features of the data with no free parameters. This is particularly impressive, as the experiments were run using different equipment, several months apart (this modelling procedure is equivalent to that presented in Baker *et al.*, 2007a).

It is also possible to fit the models directly to the data from this experiment (not shown).

Model		Two Stage	Two Stage	Twin Sum	Twin Sum
Subject		DHB	DJH	DHB	DJH
Figure		3.7A	3.7B	3.7C	3.7D
RMS error (dB)		1.54	1.77	1.27	1.71
Parameters	m	1.50	1.44	1.49	1.45
	n	-	-	1.30	1.24
	S	0.20	0.60	-	-
	p	7.50	8.18	2.60	2.02
	q	6.57	7.07	2.60	1.94
	Z	1.29	0.21	7.95	3.62
	k	0.16	0.16	0.19	0.20

Table 3.2: Parameters and RMS errors for simplex fits to the data of Experiment 2, for both models (not shown graphically).

This procedure yields somewhat better fits (see Table 3.2), but requires 6 free parameters. From either procedure, it is clear that both models give a good account of these results. This is a pleasing finding, as it shows that both models can generalise well to parts of the stimulus space other than those for which they were designed. As noted for Experiment 1 the Twin Summation model achieves marginally better fits. However, the Two Stage model still performs adequately well and cannot be discounted as a plausible architecture.

3.3 Experiment 3: contrast matching

3.3.1 Introduction

Experiments 1 and 2 showed that both the Two Stage and Twin Summation models provide a good description of discrimination performance at different pedestal levels. However, this reveals little about an observer's subjective experience of contrast. The models produce only a numerical response, which is assumed to relate to activation in the detecting mechanisms. If there is a direct relationship between this response and perceived contrast, then larger responses should appear as perceptually higher contrasts. It also follows that any combinations of contrasts sent to the two eyes which produce an equal response should look the same. This can be assessed empirically using a contrast matching task, in which the contrast of a test stimulus is adjusted until it appears the same as that of a standard stimulus.

Legge and Rubin (1981) performed such an experiment. They matched a binocularly unequal test pair to a binocularly equal standard (see also Cogan (1987) for a description of a similar experiment carried out by Birch (1979)). The magnitudes of the test contrasts was increased in ratio (see Figure 1.2), using the method of adjustment. Many ratios were used, and the experiment was performed at four different standard contrasts, and several spatial frequencies. As shown by Meese *et al.* (2006), both surviving models predict Legge

and Rubin's results qualitatively. However, a quantitative assessment is also desirable, using the same procedures and observers as above.

Experiment 3 required observers to match either a monocular, or a binocularly equal, test stimulus to a binocularly unequal standard. The standard contrasts were the same contrast pairs as used in Experiment 2, with the exception of the condition in which both left and right contrasts were -10dB. This condition was omitted, as it was subthreshold and therefore not appropriate for a matching task.

3.3.2 Methods

Apparatus and Stimuli

A VSG2/5 (Cambridge Research Systems, Ltd., UK), Clinton Monoray Monitor, and shutter goggles were used.

Procedure

A 2IFC procedure was used. In one interval the standard was shown, in the other was the test. The test stimulus was either monocular or binocularly equal, and contrast was determined by a 1-up 1-down staircase. The step size of the staircase, 1.5dB, was smaller than previously (Experiments 1 and 2 used 2dB and 3dB steps respectively), as psychometric functions for matching tend to be very steep ($\beta \simeq 5$). A smaller step size allows for a fuller sampling of the psychometric function.

The experiment was blocked by standard, with each block containing four staircases - two for the binocular, and two for the monocular match condition. In the monocular condition, the test was presented to the left eye for one staircase, and the right eye for the other, to achieve counterbalancing across the eyes. The unstimulated eye always viewed mean luminance.

Since the task was subjective, no auditory feedback was given, as there was no 'correct' choice.

Observers

As for Experiment 2, subjects DHB and DJH participated in the experiment. DHB repeated the experiment four times (160 trials per point), but as psychometric functions were steep there was little variation between sessions. DJH, therefore, repeated the experiment only once (40 trials per point) but generated very similar results.

3.3.3 Results

Figure 3.9 shows results for both subjects in both conditions. Curves are the predictions of the Two Stage model, based on the fits to Experiment 1 (Table 3.1). In Figure 3.10, the data are replotted along with equivalent predictions of the Twin Summation model. All predictions are therefore made with no free parameters.

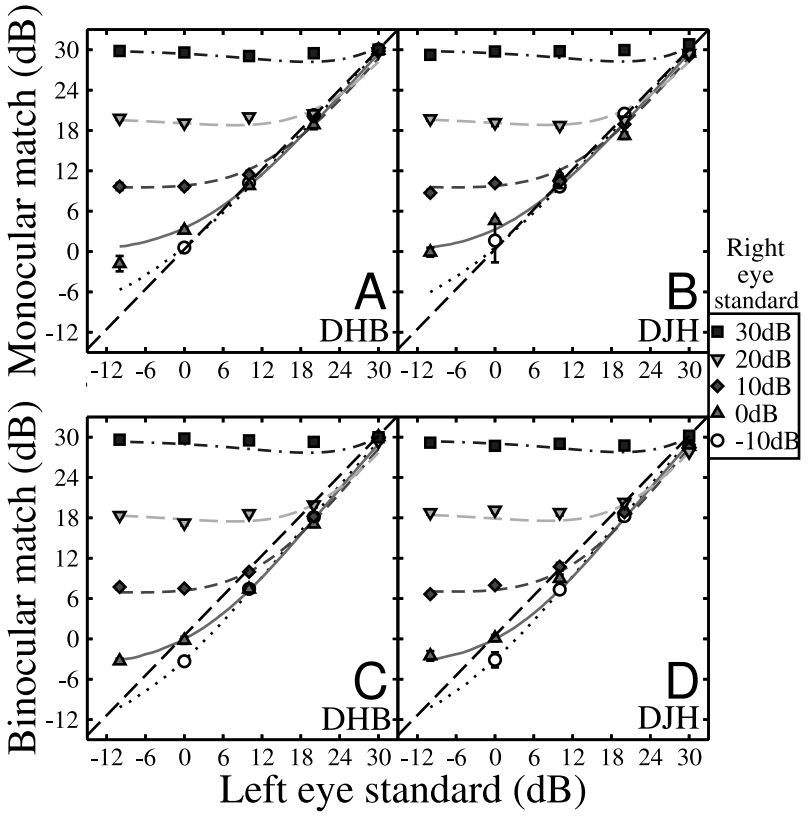

Figure 3.9: Contrast matching data for two subjects. Standards were of unequal contrasts in the two eyes, given by symbols and abscissa. Matching stimuli were either monocular (upper panels) or binocularly equal (lower panels). Curves show predictions of the Two Stage model, based on the parameters in Table 3.1. RMS errors, by panel, were A=0.88dB; B=0.98dB; C=0.97dB; D=0.78dB.

In the monocular condition, when the standard contrast in the right eye is weak (-10dB, circles), the match is almost veridical, and falls along the diagonal line of unity. As right eye standard contrast is increased, more contrast energy is required to achieve a match, and the functions prise away from the line of unity. When right eye standard contrast is high (30dB, squares), the function is almost flat, indicating that the match is largely independent of the left eye standard contrast. Both models predict this result because, at high contrasts, monocular and binocular inputs of equal magnitude produce an almost identical response, owing to the gain control features of both models. Indeed, performance is excellent for both models, with RMS errors generally below 1dB.

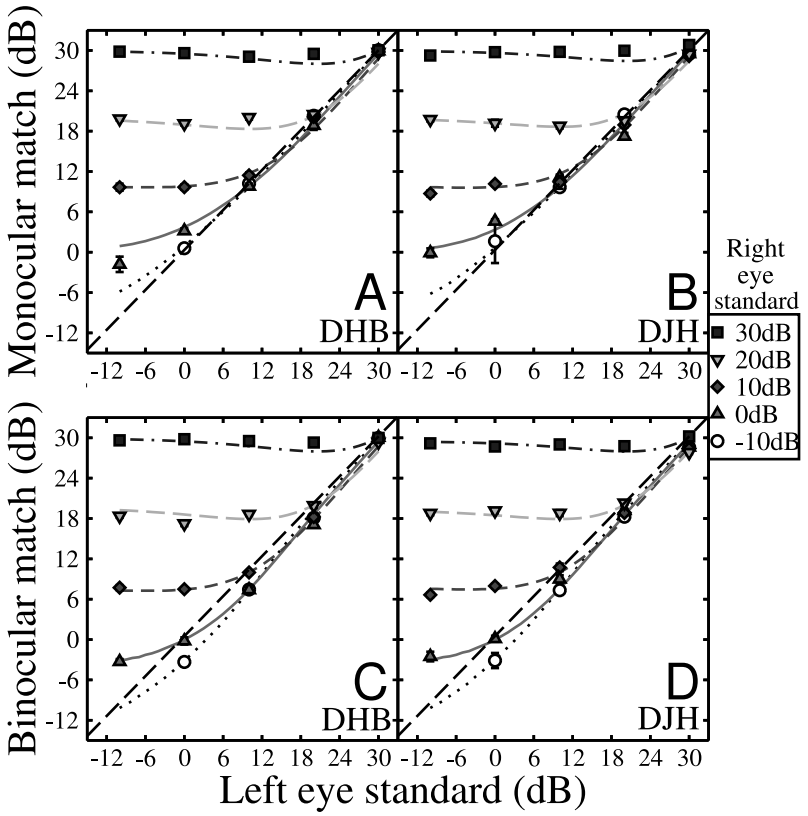

Figure 3.10: Predictions of the Twin Summation model, based on parameters in Table 3.1. Data are duplicated from Figure 3.9. RMS errors, by panel, were A=1.00dB; B=0.91dB; C=0.89dB; D=0.67dB.

3.3.4 Models of matching

It should be pointed out that some portions of the models do not influence matching performance. For example, in both models, the parameter k (see equation 3.2) is not required, as a match is obtained when the model outputs to standard and test are equal. In the Two Stage model, no part of the model after binocular summation contributes to the model's predictions. If two stimuli produce an identical result at the point immediately after binocular summation, they will produce an identical result at the output stage of the model also. Thus, only the parameters m and S are important in this task. The Twin Summation model is less straightforward, as the binocular components are not fully combined until the gain control stage, so all of the parameters contribute to the model's performance.

As demonstrated in Baker *et al.* (2007a), models such as the quadratic summation model (Legge, 1984b), or a linear summation model (Campbell and Green, 1965) are unable to predict these data. Instead, they tend to overestimate the monocular match required for a high contrast binocular standard, and underestimate the binocular match required for a high contrast monocular standard. For the linear summation model, it is obvious why this comes about, as the response to a binocular stimulus of a given contrast will always be twice that of a monocular stimulus of the same contrast. A similar situation occurs for quadratic summation, or, indeed, any nonlinear summation with an exponent >1.

Clearly, the matching data do not support such a simple front end, and require some further process, such as a gain control. In the Two Stage model, the interocular suppression at stage one reduces the effects of adding contrast in the second eye, such that at high contrasts, mon=bin, as observed experimentally. This has some similarity to the convergence of monocular and binocular dipper handles in Experiment 1. An equivalent process occurs in the Twin Summation model, caused by the 'cascade' of exponents on the numerator and denominator of the gain control.

3.3.5 Discussion

It is an important feature of any model to be able to generalise to a different task from that which it was designed for. The discrimination tasks in Experiments 1 and 2 were performance measures, whereas the matching experiment reported here was a perceptual task (Swanson, Wilson and Giese, 1984). Thus, both models have been shown to make excellent predictions of both performance and perception for contrasts of different magnitudes, presented to the two eyes. This suggests that the model response surfaces (such as those shown in Figure 3.6) are accurate in terms of both their gradient (discrimination) and their height (matching).

3.4 Experiment 4: paradoxical psychometric functions

3.4.1 Introduction

Careful inspection of Figure 3.6 will reveal that when input contrast to one eye is high, presenting contrasts to the other eye does not result in a direct increase in response. Instead, the response actually reduces slightly, before increasing again as the inputs to the two eyes become more similar. As noted in Meese *et al.* (2006), this behaviour is present in both models, and produces an interesting prediction for dichoptic contrast discrimination at high mask contrasts. Instead of the typical sigmoidal shape, the model predicts that psychometric functions under these conditions will actually drop below 50% correct in a 2AFC paradigm. In other words, an increase in contrast to one eye produces a decrease in the model response.

Figure 3.11 illustrates this behaviour for the Two Stage model. The abscissae give the mask contrast and the test contrast, which are presented to opposite eyes (dichoptic presentation). The ordinate gives the proportion of correct responses in a 2AFC task. At low mask contrasts, standard psychometric functions tending between 0.5 and 1 are produced. The functions are steep at very low mask contrast levels, become shallower in the region of dichoptic facilitation, as also shown in Figure 3.5 (and Bird *et al.*, 2002). At high mask contrasts, increasing test contrast causes the predicted performance to fall to well below chance levels, before increasing rapidly.

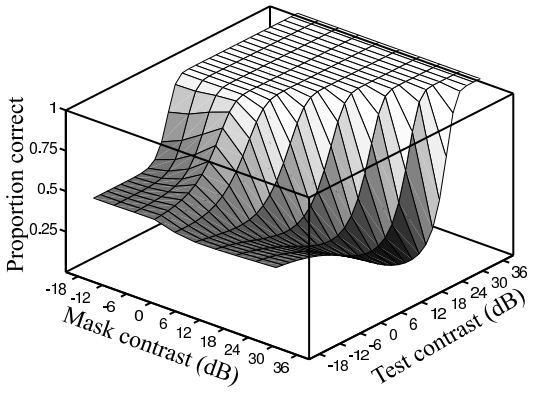

Figure 3.11: Psychometric surface, with dichoptic mask and test contrasts given on the abscissae, and predicted proportion of correct responses in a 2AFC task on the ordinate. At low mask contrasts, psychometric functions have a typical sigmoidal shape, between 0.5 and 1. For higher mask contrasts, performance drops below chance levels.

This is, therefore, an instance of negative d' (Green and Swets, 1966), which results in a non-monotonic psychometric function. Since there is a widely held assumption that the psychometric function increases monotonically with contrast (i.e. García-Pérez and Alcalá-Quintana, 2007), this highly unusual, indeed paradoxical, prediction obviously warrants empirical exploration.

3.4.2 Methods

The apparatus, stimuli, experimental software and observers were the same as for Experiment 1.

Procedure

The method of constant stimuli was used, to ensure that psychometric functions were sampled adequately across their full extent. Four psychometric functions were measured, with mask contrasts of 0%, 5dB, 20dB and 30dB. These levels were selected to test the prediction that at low mask contrasts, functions are sigmoidal, but that at high mask contrasts they are not. Eleven test contrasts were selected, in steps of 2dB, around the threshold values recorded in Experiment 1.

Presentation was blocked by mask contrast, and all test contrasts interleaved within a block. 40 trials per test contrast level were gathered on each repetition, and each observer completed 5 repetitions. This means that each point in Figure 3.12 is the result of 200 trials, and that each observer completed 8,800 trials overall ($200 \times 11 \times 4$).

As for Experiment 3, no auditory feedback was given. However, in this experiment, this was done to prevent confusion for the observers. If the models are correct, in the paradoxical region, the mask interval should appear to have a higher contrast than the mask+test interval. As observers were instructed to select the interval with the highest apparent contrast, and auditory feedback indicates correct selection of the test interval, in the paradoxical case, correctly responding to the higher contrast would result in feedback indicating an incorrect response.

3.4.3 Results

The results are shown in Figure 3.12. Curves are predictions of the Two Stage model (upper) and the Twin Summation model (lower), based on the parameters in Table 3.1 for each subject. The results for the two low mask contrast functions are predicted well. Here, the change in psychometric slope and reduction of threshold for the 5dB mask contrast, also noted in Experiment 1, are evident in both model and data. It is reassuring to have replicated these two findings using a denser sampling of the psychometric function than that afforded by a staircase procedure.

The paradoxical regions at higher mask contrasts are not consistent with the model predictions. All three observers show a smaller 'trough' region than predicted by the model. However, although it is clear that the predictions are not quantitatively accurate, there

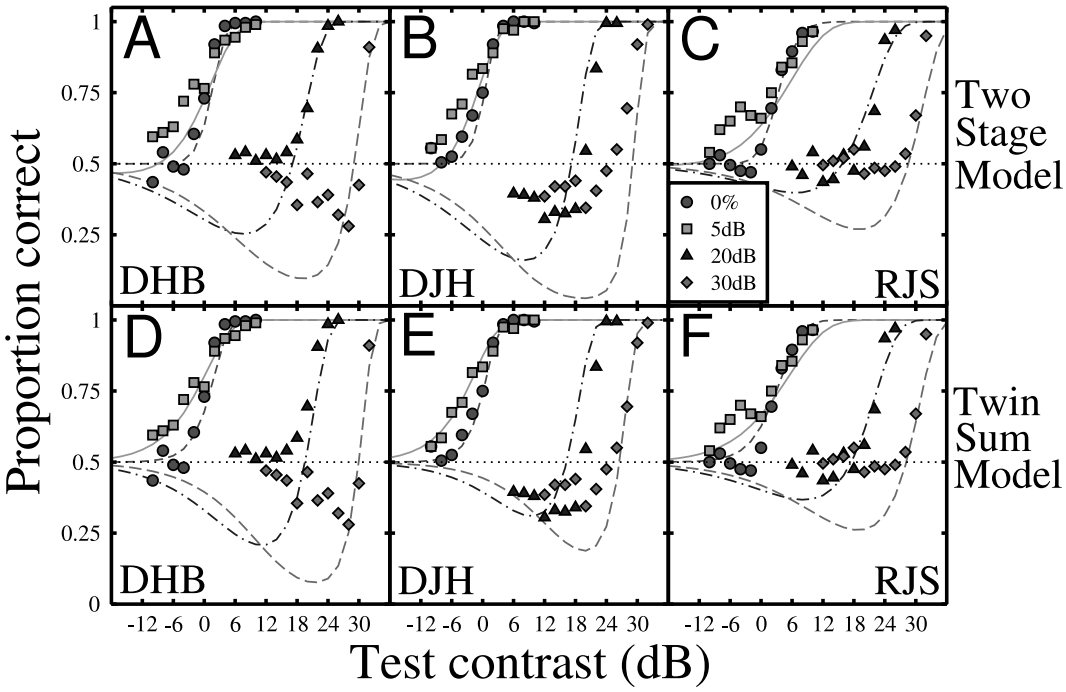

Figure 3.12: Paradoxical dichoptic psychometric functions for three observers. Data points represent 200 trials each in a 2AFC, method of constant stimuli experiment. For this large number of trials, binomial error bars are smaller than symbols in all cases, and so are not shown. Curves give predictions of the Two Stage (upper) and Twin Summation (lower) models, based on the parameters in Table 3.1.

is some empirical evidence for the paradoxical effect. Subjects DHB and DJH both show 'trough' regions at 30dB mask contrast, and DJH also shows one for the 20dB mask. Subject RJS shows no sign of the paradox. It is interesting to note that the model predictions show the correct trend: the parameters derived for RJS from Experiment 1 predict a smaller paradoxical effect than do the parameters for DHB or DJH.

3.4.4 Discussion

Psychometric functions were recorded for dichoptic pedestal masking at four mask contrasts. As predicted by two models of binocular contrast vision, at high mask contrasts the psychometric functions were non-monotonic for two observers. In the model this is due to a highly nonlinear relationship between test contrast and model response (d'), and it seems reasonable to assume a similar cause for human subjects. This surprising behaviour is apparently unique; I am unaware of any similar reports in the literature

Although the models do predict the effect, there is definite room for improvement. Several modifications and fitting procedures were attempted, all of which were ultimately unsuccessful. These included adjusting the weight of dichoptic masking, comparing results from left and right eyes, a ratio based model inspired by that of de Weert and Levelt (1974), fitting the

models to the entire 'psychometric surface' (Wichmann, 1999), and introducing stochastic elements. Typically, altering the models to produce accurate predictions for the Experiment 4 data severely worsened performance for other tasks, such as the dipper functions of Experiment 1. Attempts to understand the paradoxical behaviour are ongoing.

3.5 Comparison of two models

Given the substantial differences in the architecture of the Two Stage model and the Twin Summation model (Figure 3.3) , it is surprising that they are so similar. The equations which determine the model responses are not mathematically equivalent, and yet they appear to behave in the same manner over the range of input values which can be tested empirically.

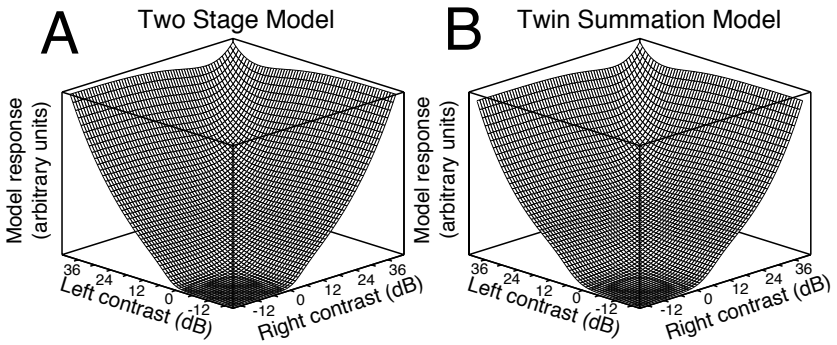

Figure 3.13: Response surfaces for the Two Stage model (A) and the Twin Summation model (B). The surfaces were generated using parameters from Meese *et al.* (2006), which were: Two Stage model, p=7.99, q=6.59, Z=0.077, m=1.28, S=0.985; Twin Summation model, p=2.47, q=2.4, Z=7.06, m=1.43, n=1.28. The absolute magnitudes of the responses differ, however these have no bearing on the predictive power of either model, so each surface is scaled to its maximum.

This is also apparent from a comparison of the model response surfaces, generated using parameters from Meese *et al.* (2006). Visual inspection of such surfaces, shown in Figure 3.13, shows them to be almost indistinguishable. Further quantitative analyses, including fitting the response surface of one model to that of the other, revealed no substantial differences. Attempts are made in the following chapter to differentiate between the models on other grounds.

3.6 Chapter summary

Four experiments have explored the suppressive and facilitatory effects which occur within individual channels in the binocular visual system. Dichoptic masking was found to be very strong at high pedestal contrasts, but dichoptic facilitation was observed at lower pedestal contrasts. Binocular summation at threshold was $>\sqrt{2}$ for all subjects. Monocular and

binocular contrast discrimination functions were similar to those reported previously in the literature. Four models were assessed, two of which gave an adequate description of the data. The models were further tested on different regions of the stimulus space, and also using a matching task. In both cases, they performed well with no free parameters. Finally, a model prediction was tested that non-monotonic psychometric functions occur at high dichoptic mask contrasts. Non-monotonic functions were found, but these were less extreme than predicted by either surviving model.

These experiments provide a comprehensive exploration of binocular within-channel processes. It is clear that both models are excellent predictors of discrimination performance across the entire range of displayable stimulus contrasts. Importantly, the models can also predict accurately the perception of different contrast levels, as demonstrated with a contrast matching task. This ability to relate perception and performance is an important feature of any model (Swanson *et al.*, 1984) and gives much credibility to those discussed here. Given that performance using within-channel stimuli has now been adequately characterised, the following chapter explores interactions between different channels.

Chapter 4

Initial investigations into interocular cross-channel masking

4.1 Chapter Introduction

This chapter explores an observer's ability to detect a grating in the presence of a mask which differs substantially in its spatial properties. As explained in Chapter 1, although these conditions have been explored extensively, both psychophysically (Campbell and Kulikowski, 1966; Foley, 1994; Foley and Chen, 1997; Meese and Holmes, 2002; Meese, 2004; Petrov, Carandini and McKee, 2005; Meese and Holmes, 2007), and neurophysiologically (Morrone *et al.*, 1982; Bonds, 1989; Nelson, 1991; Freeman *et al.*, 2002), most work to date has used only binocular stimuli. Where dichoptic cross-channel masking has been investigated, it has usually been explored in isolation, and not compared to the monocular or binocular cases (i.e. Bacon, 1976, Levi *et al.*, 1980, though see Legge, 1979; Meese and Hess, 2004).

This chapter contains a systematic exploration of interocular cross-channel masking, using several ocular combinations of mask and test. The intention was to provide a complete picture of cross-channel masking which would allow for detailed quantitative modelling, and extend knowledge of cross-channel binocular interactions. Two experiments are reported, followed by a modelling section.

4.2 Experiment 5: Cross-channel masking in five ocular configurations

4.2.1 Introduction

The first aim was to perform a detailed exploration of interocular cross-channel masking. There are several possible ocular combinations of mask and test, all of which were investigated. The monocular, binocular, and dichoptic conditions were as described in Chapter 3, except that the mask always had different spatial properties from the test. Two new conditions were also tested. In the half binocular condition, the mask was presented to both eyes, and the test to only one eye (see Meese *et al.*, 2006; Meese and Hess, 2005). In the converse situation, the mask was presented to one eye, and the test to both. This is referred to as the mon-mask condition. All five conditions are shown diagramatically in Figure 4.1.

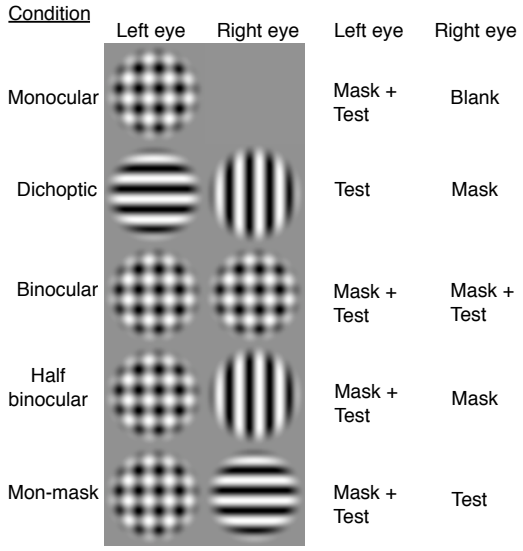

Figure 4.1: Diagram of stimulus configurations for cross-channel experiments. Examples are shown for vertical mask and horizontal test of the same spatial frequency. Conditions were always counterbalanced across eye of presentation.

Following Meese (2004) and Meese and Hess (2004), two different masks were used, which differed in their spatial properties. One was orthogonal to the test, and had the same spatial frequency. The other was obliquely oriented and had a spatial frequency three times that of the test. Masking was measured in all five ocular configurations, with both mask types.

4.2.2 Methods

Apparatus and Stimuli

Stimuli were presented on a Clinton Monoray monitor using a ViSaGe stimulus generator. All stimuli were viewed through a mirror stereoscope, to allow presentation of different stimuli to the two eyes. To allow for differences in interpupillary distance, the stereoscope was calibrated for each observer using the subjective calibration procedure described in Appendix A. After calibration, binocular fusion of the two images was effortless. Additionally, all stimuli were presented in the centre of a circular aperture, 9° in diameter. The display was set to mean luminance within the aperture, and was dark outside it. This method produced a very strong cue to fusion, which allowed the observer's visual system to compensate for any slight errors in mirror calibration.

All stimuli were generated as described in Chapter 2. Test stimuli were always horizontal, and at 1cpd. Mask stimuli were either vertical at 1cpd, or oblique (-45°) and at 3cpd. The spatial window was identical for all stimuli, and was as described previously (5° diameter, 3° plateau). Eight mask contrast levels were tested: 0 to 30dB in steps of 6dB, as well as 33dB and 0%. The 33dB condition was included as it was the highest dB integer mask contrast displayable using this equipment.

Procedure

Subjects were seated in a darkened room, with their head in a support to which the stereoscope was attached. The optical viewing distance was 57cm. Stimuli were blocked by mask contrast, with the five ocular conditions interleaved within each block. For each condition, two independent 3-down 1-up staircases were run to counterbalance across eye of presentation (in the binocular condition, these were identical). As previously, the data were collapsed across eye of presentation and session, and thresholds at the 75% correct point were estimated using probit analysis (Finney, 1971). Blocks were run in random order, and each subject completed four replications of the experimental conditions in which they participated.

As well as the two mask types, two stimulus durations were also explored, 50 and 200ms. Initially, these durations were chosen because of concerns that binocular rivalry may affect cross-oriented stimuli presented for more than around 150ms (Wolfe, 1983). While this did not appear to be an issue during the pilot study, the two durations did produce thought-provokingly different results, so they remained part of the experimental design.

Observers

Three observers participated in the experiment. All were emmetropic, with no anomalies of binocular or stereo vision, and were experienced psychophysical observers. Subject DHB (Interpupillary distance (IP)=6cm) completed the experiment with both mask types, at both durations. Subject TAY (IP=6.5cm) ran the experiment with the orthogonal mask at 200ms, and subject SAW (IP=7.5cm) viewed the oblique mask at 200ms.

4.2.3 Results

The results for the orthogonal mask are shown in Figure 4.2, and those for the oblique mask in Figure 4.3. In each figure, the upper panels contain results for the monocular, dichoptic and half binocular conditions, and the lower panels contain results for the binocular and mon-mask conditions. In all conditions, the level of masking increased with mask contrast, consistent with previous studies.

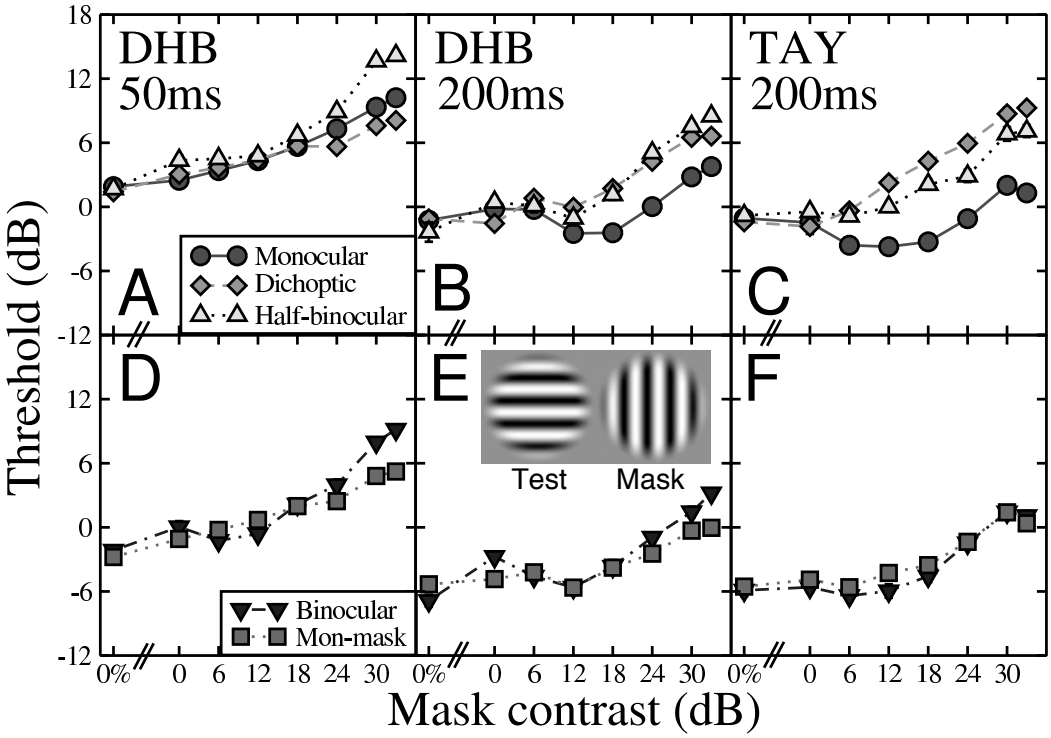

Figure 4.2: Interocular cross-channel masking results using an orthogonal mask stimulus (1cpd). Data are shown for monocular, dichoptic and half binocular conditions in the upper panels, and for binocular and mon-mask conditions in the lower panels. Observer initials and duration are given for each pair of panels. Error bars show the standard error of the probit fit, and in most cases are smaller than symbols. Binocular summation ratios at threshold were 1.61 and 1.68 for DHB (50 and 200ms respectively) and 1.70 for TAY.

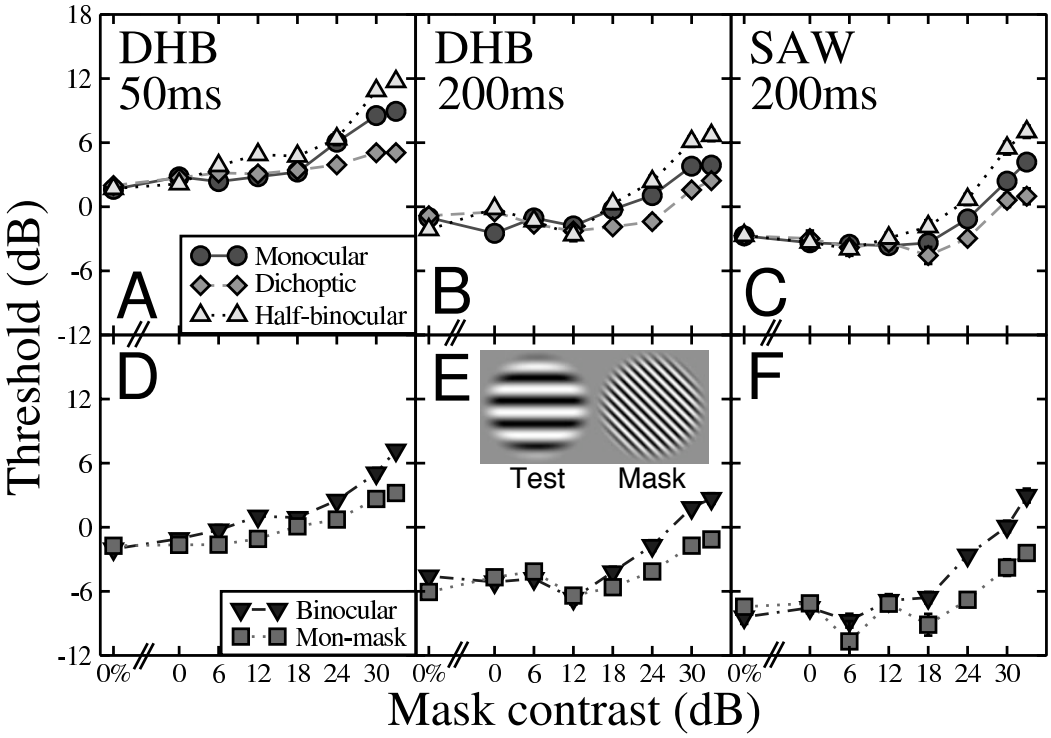

Figure 4.3: Interocular cross-channel masking results using an oblique mask stimulus (3cpd). Format is the same as for Figure 4.2. Binocular summation ratios at threshold were 1.52 and 1.58 for DHB (50 and 200ms respectively) and 1.82 for SAW.

Monocular and dichoptic masking

In the monocular and dichoptic conditions, mask and test are each presented to one eye only. The only difference is whether they are presented to the same or different eyes. Given that mask energy is equal for these two conditions, one might expect to observe identical levels of masking. However, the masking functions for these conditions clearly do not superimpose. Monocular masking was generally strongest (circles are above diamonds in Figs 4.2A and 4.3A-C), apart from in the orthogonal condition at 200ms (Fig 4.2B, C), where dichoptic masking is clearly stronger for both observers (circles below diamonds).

This finding suggests that there may be separate processes that give rise to dichoptic and monocular cross-channel masking. The differences appear to depend on both mask type (oblique or orthogonal) and stimulus duration (50 or 200ms), it seems likely that these mechanisms have different spatial and temporal tuning properties. Experiment 6 explores the effect of stimulus duration systematically, and the tuning of monocular and dichoptic cross-channel masking is reported in Chapters 5 and 6.

Other conditions

The half binocular condition is interesting, as the mask energy across the eyes is greater than that in either the monocular or dichoptic conditions, while the test remains the same. It seems reasonable to expect masking to be greatest in this condition. In general this is true (triangles are highest in Figures 4.2A and 4.3A-C), however for the orthogonal mask at 200ms, there are regions in which the half binocular function sits below the dichoptic function (triangles below diamonds in Figure 4.2C). This unexpected finding, termed the half binocular paradox, will be a challenge for models of masking to account for.

In a similar vein, the binocular condition generally elicits greater masking than the mon-mask condition, as it has greater mask energy (Figures 4.2D, E and 4.3D-F). The exception to this rule is shown in Figure 4.2F. Here masking appears to be equal across the two conditions. This may be caused by the same process which produced the half binocular paradox described above, as it occurs for the observer who shows the paradox most strongly (TAY). Although the binocular condition has a mask in each eye, it does not produce substantially greater masking than the mon-mask condition, in which the mask is shown to only one eye.

Binocular summation of test signals

The binocular summation ratio at threshold for the whole data set was 1.65 (4.36dB), consistent with the findings reported in Chapters 1 and 3. In addition, the binocular and half binocular conditions can be compared directly to give us an estimate of the amount of binocular summation (see also Experiment III of Meese *et al.* (2006)) at all mask contrast levels. These functions are virtually parallel for all subjects in all conditions (see Figure 4.9 below). Figure 4.4 shows the summation ratio ($\frac{thresh_{hbin}}{thresh_{bin}}$) at all mask contrast levels. The stars give the average across all six functions. Aside from a small number of outliers, the ratios lie between $\sqrt{2}$ and 2, consistent with Chapters 1 and 3, as well as Meese *et al.* (2006). The mean summation ratio across all mask contrasts was 1.71 (4.66dB), which also compares favourably with earlier conclusions. A similar comparison (not shown) can be made between the monocular and mon-mask conditions. This yields similar results, although summation (ratio=1.57) is reduced somewhat, probably because of the presence of monocular facilitation (see next section).

Monocular facilitation

There is a small amount of facilitation (∼3dB) in the monocular condition for the orthogonal mask at 200ms (Fig 4.2B, C). Cross-channel facilitation has been reported widely in

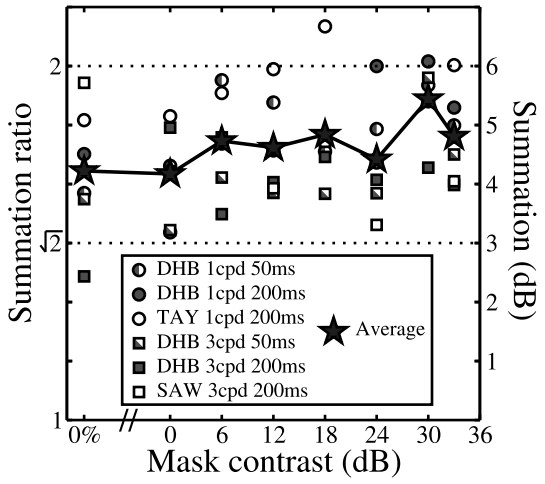

Figure 4.4: Binocular summation assessed across the entire masking function. Ratios are calculated by dividing the threshold for the half binocular condition by that for the binocular condition. The functions show no consistent upward or downward trend, but vary randomly between $\sqrt{2}$ and 2. Stars are the average across all six functions.

the literature (Yu, Klein and Levi, 2002; Meese and Holmes, 2006, 2007; Meese, Summers, Holmes and Wallis, 2007b) and is generally of a similar magnitude to that observed here. It is possible for facilitation to reduce the slope of the psychometric function, as observed for within-channel stimuli in Chapter 3 (see also Bird *et al.*, 2002). If this were the case, it would suggest that facilitation was a within-channel effect, caused by the mask partially stimulating the detecting mechanism. However, for both subjects who showed the facilitation, the psychometric slopes did not vary substantially from those observed at detection threshold ($\beta \simeq 4$), arguing against a within-channel account. It also seems unlikely that the facilitation was caused by reduction of spatial uncertainty (Petrov *et al.*, 2006), as it was not observed either in the dichoptic orthogonal condition, nor in any condition using the oblique mask stimuli. These masks should provide the same location information as the monocular orthogonal mask, and so reduce spatial uncertainty by the same amount.

Different temporal dependencies?

One particularly unexpected finding was that for the orthogonal condition, the ordering of monocular and dichoptic functions changed as the stimulus duration was increased from 50 to 200ms (Fig 4.2A and B). This suggests that the mechanisms which determine suppression in these two configurations may have different temporal dependencies. Experiment 6 was devised to explore this possibility.

4.3 Experiment 6: Cross-channel masking at varying stimulus durations

4.3.1 Methods

The equipment, stimuli and observers were the same as for Experiment 5. Masking was measured in the monocular and dichoptic conditions only, with separate staircases to control for eye of presentation. Stimuli were blocked by duration (25, 50, 100, 200 and 400ms), with all conditions interleaved in each block. The mask contrast was fixed at 33dB. Baseline detection thresholds (mask contrast of 0%) were also recorded at each duration, but were measured separately from the masked conditions. After the main experiment had been carried out, the half binocular condition was also run under the same conditions. All subjects repeated the experiment four times, and with each mask type.

4.3.2 Results

Results are shown in Figure 4.5. Upper panels give performance with the orthogonal mask, lower panels for the oblique mask. The baseline thresholds (which are duplicated for each subject in upper and lower panels) decreased with duration, and reached asymptote at around 100ms, consistent with previous studies (Nachmias, 1967; Tolhurst, 1975a; Legge, 1978; Luntinen, Rovamo and Nasanen, 1995).

Once again, it is clear that the monocular and dichoptic functions do not superimpose. What is more, they appear to show opposite effects of duration. Relative to the baseline, the dichoptic functions increase with stimulus duration, whereas the monocular functions are either flat, or decrease with duration. This is particularly apparent in Figure 4.6, which shows the data replotted as threshold elevation relative to baseline. Furthermore, this opposite direction of effect causes the monocular and dichoptic functions to cross over for one subject (DHB, see Figures 4.5A and 4.6A), consistent with the change in magnitudes observed in Experiment 5 for this subject.

For the half binocular condition, the paradoxical reduction in masking relative to the dichoptic condition is seen once again, this time in all subjects (triangles below diamonds in panels A-C of Figure 4.5). However, it appears to occur only in the regions where dichoptic masking is much stronger than monocular masking (Fig. 4.5A-C, E). When monocular masking is greatest, even by a similar amount, the paradox is not observed (Fig. 4.5D & F).

As will be discussed below, the pattern of monocular and dichoptic thresholds indicate that there are two distinct routes to suppression, one monocular and one dichoptic. The

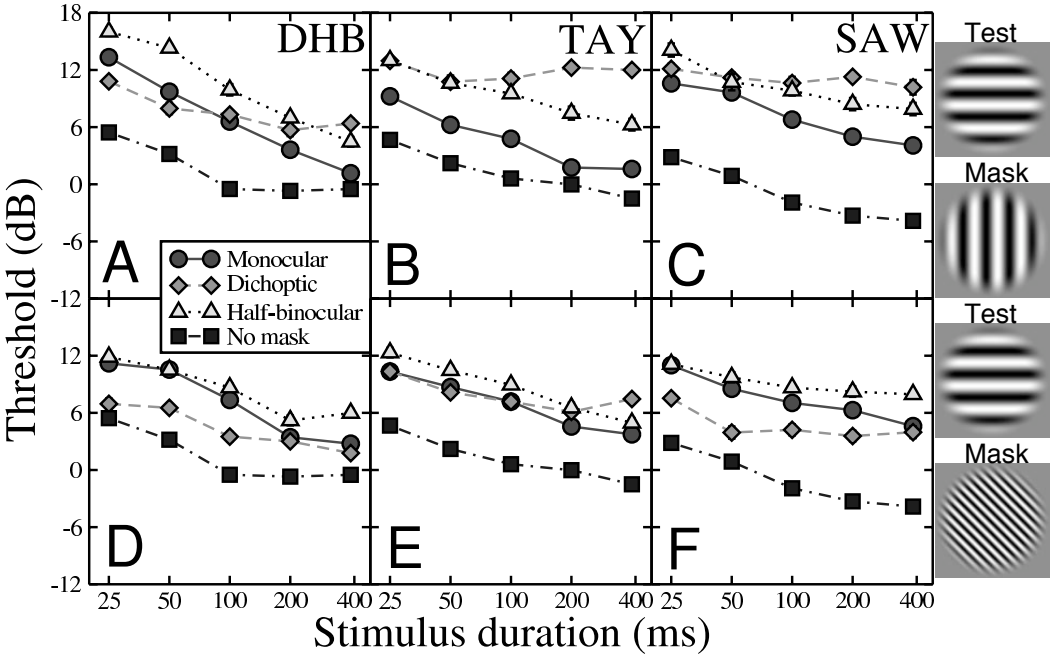

Figure 4.5: Cross-channel masking at varying stimulus durations. Upper panels show results for the orthogonal (1cpd) mask, and lower panels show results for the oblique (3cpd) mask. Filled squares are the baseline detection thresholds (0% mask contrast), and other symbols give thresholds with a high contrast (33dB) mask, presented in the ocular configurations shown in Figure 4.1 and given by the legend.

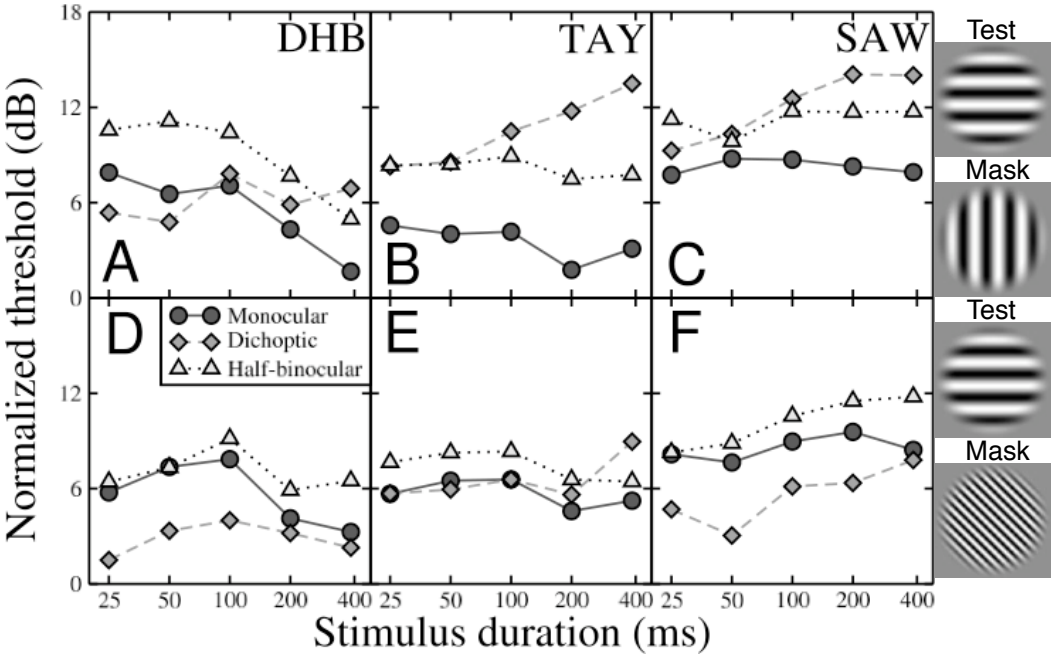

Figure 4.6: Cross-channel masking at varying stimulus durations, normalized to baseline detection thresholds at each duration. Panel ordering and symbols are the same as for Figure 4.5.

half-binocular paradox could occur if suppression from the dichoptic route is reduced by the introduction of an additional mask component. This would decrease the level of suppression, relative to the dichoptic condition alone. Since the half-binocular thresholds are never below the monocular thresholds, the monocular channel is presumably immune to this effect.

4.4 Summary

Three important results have been identified by these two experiments:

- monocular and dichoptic masking have different magnitudes;
- they show a different pattern of results with stimulus duration;
- in the half binocular condition, the extra mask can reduce the level of masking relative to the dichoptic condition.

These findings all indicate the existence of two separate mechanisms, which govern monocular and dichoptic masking. This is consistent with recent reports from the physiology literature (Li *et al.*, 2005; Sengpiel and Vorobyov, 2005), showing that monocular and dichoptic masking display the characteristics of precortical and cortical phenomena, respectively. Further discussion of the relevant physiological findings is deferred to Chapter 5.

4.5 Extending the Two Stage Model

4.5.1 Basic model

Within the architecture of the Two Stage model, there are numerous points at which cross-channel masking could potentially impact. Initially, masking is assumed to occur only at the gain control stages. Heeger (1992) proposed the concept of a suppressive 'gain pool', which would act divisively on a detecting neurone, and this idea is adopted for the present model, as has been done previously for binocular masking (i.e. Foley, 1994; Meese and Holmes, 2002; Holmes and Meese, 2004).

If masking occurred solely at a late stage, after binocular summation, then monocular and dichoptic masking should be identical (Figure 4.7A). This is because by the time masking occurs, eye of origin information has been lost, and a mask will have the same effect, regardless of which eye it was presented to. From the data presented in Figures 4.2 - 4.5, this is clearly not the case, so models with a single, late site of masking can be ruled out.

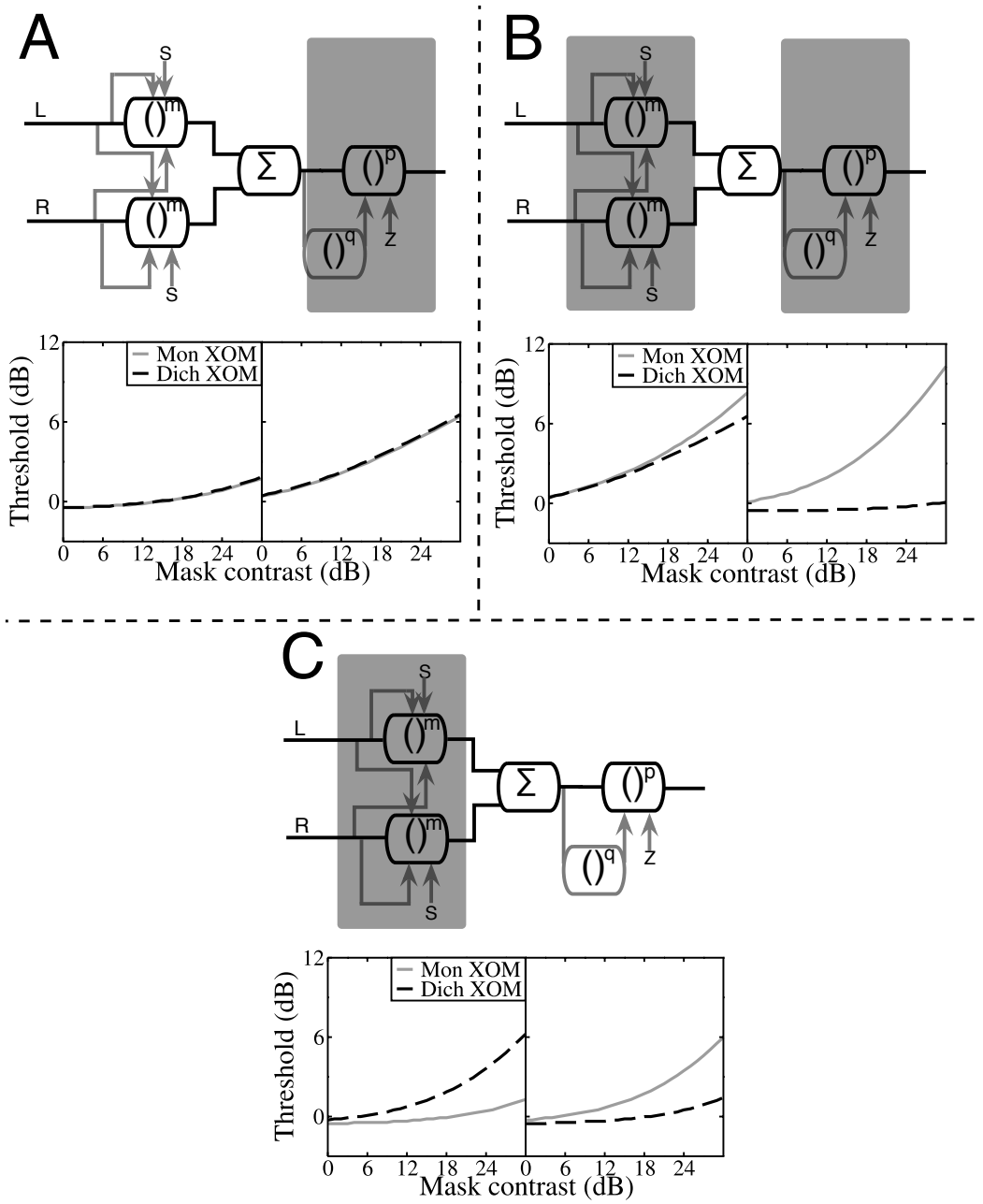

Figure 4.7: Illustration of how the point of impact of monocular and dichoptic suppression constrains the relative magnitudes of the masking functions. In all sections, the shaded region gives the locus at which cross-channel suppression impacts. For each configuration, two example graphs are provided, illustrating the behaviours available using different parameter values. For late suppression (A), monocular and dichoptic masking functions superimpose. When one type of suppression also impacts before stage one (B - here, monocular), the masking from that route must be greater. If both forms of suppression occur at stage one, the magnitudes of monocular and dichoptic masking are independent (C), and either can be the greater.

A further possibility is that masking impacts before binocular summation in one case, with a late, post-summation site producing masking for the other ocular configuration. This arrangement would require that whichever process occurs before binocular summation (be it monocular or dichoptic) will necessarily result in the strongest masking (Figure 4.7B illustrates early monocular masking). This is because the early masking is augmented by further masking at the later stage, so for an equivalent mask contrast will always be stronger than masking from the late stage alone. Since monocular and dichoptic masking appear to have independently varying magnitudes (see Fig. 4.5A for an example of them crossing over) arrangements of this nature are clearly not sufficient.

One configuration which does allow for the breadth of results observed empirically, is if both types of masking occur prior to binocular summation (Baker and Meese, 2006a,b; Baker *et al.*, 2007c). By assigning different weights to monocular and dichoptic masking, their magnitudes can vary independently (Figure 4.7C). The model favoured by Baker *et al.* (2007c) was a simplified version of the Two Stage Model, in which masking was represented in this manner. An obvious starting point is to extend the full Two Stage Model to incorporate cross-channel masking at stage one. The full version of the model will then be able to predict performance for both within- and cross-channel mask stimuli.

The equation for stage one now becomes,

$$stage1 = \frac{{C_L}^m}{S + C_L + C_R + \omega_M X_L + \omega_D X_R} + \frac{{C_R}^m}{S + C_R + C_L + \omega_M X_R + \omega_D X_L}, \qquad (4.1)$$

where X_L and X_R are mask contrasts to the left and right eyes, and ω_M and ω_D are the weights of monocular and dichoptic masking for a given mask. Other terms, and stage two of the model, are the same as in Chapter 3. This is referred to as Model 1.

Most of the model parameters do not play a major role in describing the present set of results. For this reason, five of the parameters are fixed at the values from Experiment II of Meese *et al.* (2006) (given here in the caption to Table 4.1). The only parameters allowed to vary are ω_M and ω_D, the weights of cross-channel masking, and also k, to allow baseline sensitivities to vary across observers and stimulus durations. The model is fit to the monocular and dichoptic masking functions only, and then used to predict performance for the other three conditions.

The best fits of Model 1 to the monocular and dichoptic data of Experiment 5 are shown in Figure 4.8, with parameter values and error statistics given in Table 4.1. The fits are generally good (RMS errors $\simeq$1dB), with the model able to capture differences in the order of masking functions by varying the weight parameters. The worst two fits are

Subject	Condition	RMS error (dB) Mon & Dich	RMS error (dB) All functions	AIC score	k	ω_M	ω_D
DHB	1cpd 50ms	0.86	1.07	8.71	0.235	0.058	0.034
DHB	1cpd 200ms	1.08	1.18	12.62	0.023	0.020	0.063
TAY	1cpd 200ms	1.36	2.08	35.30	0.012	0.014	0.147
DHB	3cpd 50ms	0.65	0.78	-3.94	0.194	0.043	0.013
DHB	3cpd 200ms	0.79	0.99	5.60	0.016	0.031	0.016
SAW	3cpd 200ms	0.84	1.11	10.17	0.003	0.043	0.020

Table 4.1: Best fitting parameters and error statistics for Model 1 (three free parameters). Fits to the monocular and dichoptic data are shown in Figure 4.8, with RMS errors given in the third column. Predictions for the other three conditions are shown in Figure 4.9, with RMS errors to the full data set (five functions) in the fourth column. AIC scores are calculated across all five functions. Model parameters other than those given were fixed at the values from Experiment II of Meese *et al.* (2006): $p = 7.99, q = 6.59, Z = 0.076, m = 1.28, S = 0.98$.

for subjects DHB and TAY for the 200ms orthogonal mask. This is the condition in which monocular facilitation is observed. The model does not capture this trend, hence the poorer fit. However, in Section 4.5.2 the model is extended to encompass facilitation.

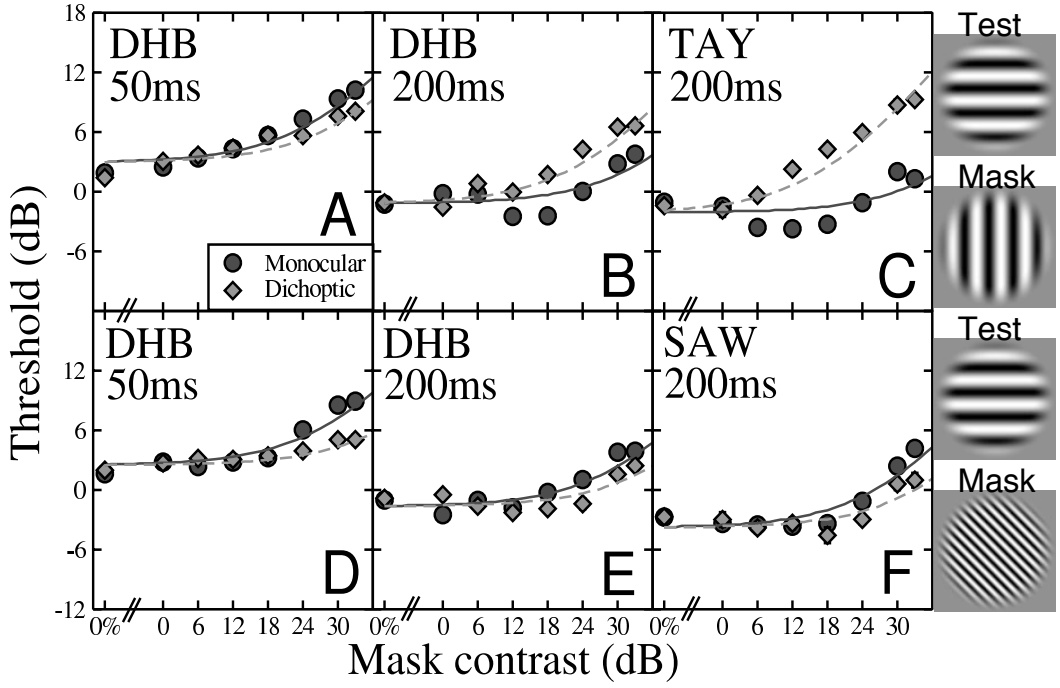

Figure 4.8: Best fit of Model 1 to the monocular and dichoptic data of Experiment 5. Upper panels are for the orthogonal mask, and lower panels for the oblique mask. Parameters and error statistics are given in Table 4.1.

The parameters estimated from fitting to the monocular and dichoptic data can be used to predict performance in the remaining three conditions with no free parameters. These predictions are shown in Figure 4.9, and RMS errors to all five functions are also given in Table 4.1. In general, the predictions are good, with the model capturing most of the separations between functions. The most notable exception is again for subject TAY. Here,

the model predicts greater masking in both the half-binocular and the binocular conditions than is observed experimentally. This is due to the half-binocular paradox described above, which Model 1 is not able to predict. Although this is only seen clearly here for one observer, the results of Experiment 6 show the paradox for all three subjects. In the following section, the model is extended to encompass the half-binocular paradox, and monocular facilitation.

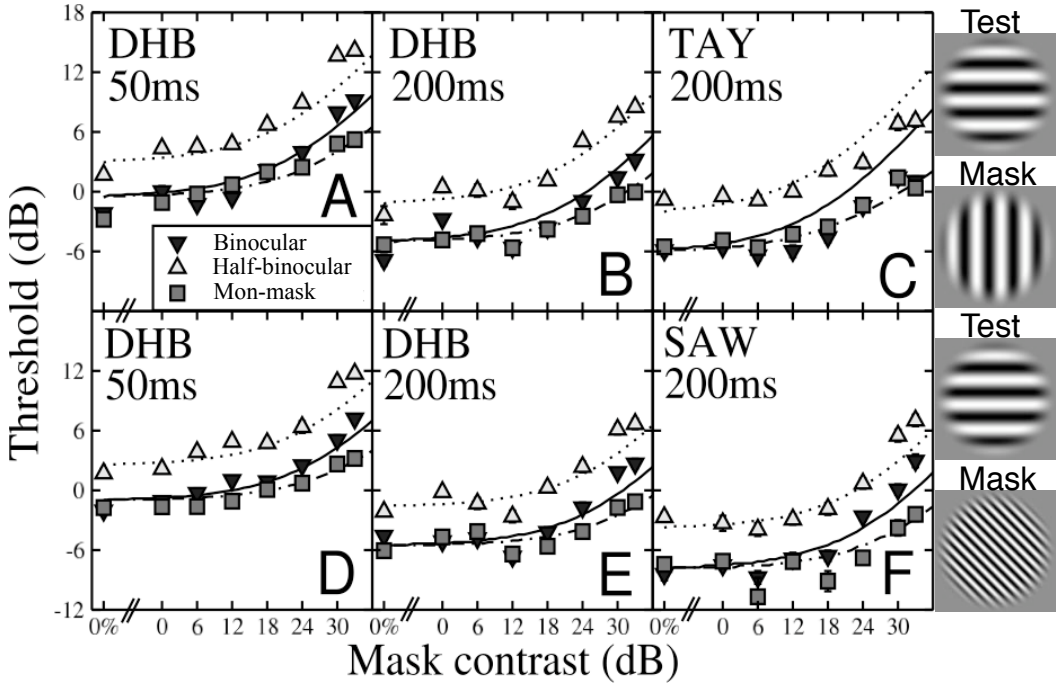

Figure 4.9: Predictions of Model 1 for the binocular, half-binocular and mon-mask conditions. Here, all parameters were fixed to those derived by fitting the monocular and dichoptic data, shown in Figure 4.8 and Table 4.1. Panel layout is the same as for Figure 4.8.

4.5.2 Describing subtleties of the data set: Models 2 & 3

In section 4.2.3, a surprising result was discussed for the half-binocular condition. Relative to the dichoptic condition, mask energy has increased (see Figure 4.1), but for some subjects the magnitude of masking decreases. This half binocular paradox does not occur in Model 1 because adding mask contrast in either eye will always enlarge the denominator term, and thus increase the amount of masking. In order to reduce masking, some additional process must occur to decrease the level of masking when two masks are present.

A similar phenomenon to the half-binocular paradox is observed using within-channel stimuli, and is predicted by the standard version of the Two Stage model (Meese *et al.*, 2006). Figure 4.10A shows the model's predictions for within-channel (pedestal) masks in dichoptic, monocular and half binocular configurations. Here, the half binocular condition displays a level of masking in between that for monocular and dichoptic masks alone, despite an overall

increase of mask energy. This behaviour is consistent with empirical findings reported in Chapter 3 and by Meese *et al.* (2006). In the model, the paradox occurs because, relative to the dichoptic condition, the additional mask both facilitates detection (the pedestal effect), and also reduces the impact of the dichoptic mask (via interocular suppression).

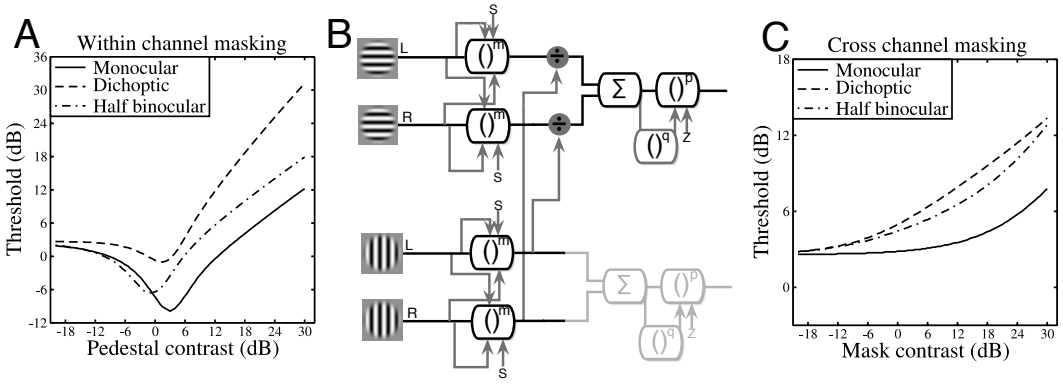

Figure 4.10: Half binocular model and predictions. A) depicts the performance of the standard Two Stage model for monocular, dichoptic and half binocular pedestal masking. The half binocular masking function falls between the other two functions, despite there being greater mask energy in that condition. B) shows the architecture of a model designed to produce the half binocular paradox for cross-channel stimuli (Model 2). Mask and test are processed in separate channels until after stage one, when they interact dichoptically (monocular cross-channel masking is not shown). C) gives predictions of this model for cross-channel stimuli in the three conditions. With appropriate parameters, the half binocular function lies between those for monocular and dichoptic.

For cross-channel masks, the 'indirect' masking effect caused by summation of mask and test channels does not occur (see Chapter 6 for further details of this). Also, facilitation does not appear to be substantial enough to produce the paradox, which occurs at contrasts where the additional (monocular) mask causes threshold elevation (i.e. Figure 4.5). However, it is possible to produce the effect if the dichoptic signal is subject to interocular suppression within its own detecting mechanism before it impacts the test channel (Fig. 4.10B). This means that when dichoptic masking is strong, the addition of an extra (monocular) mask will decrease the level of dichoptic masking by a greater amount than it causes masking itself (see Fig. 4.10C).

The first stage, which occurs in both mask and test channels, is now,

$$stage1_L = \frac{{C_L}^m}{S + C_L + C_R + \omega_M X_L}, \tag{4.2}$$

where C refers to the contrast of gratings to which the mechanism is most sensitive, and X refers to the contrast of gratings which do not excite the mechanism. Since this stage includes within-channel dichoptic masking (see Chapter 3), the output for a dichoptic mask will be reduced when an extra mask component is added in the half binocular condition. Dichoptic cross-channel masking then occurs after stage one, and is divisive. Dichoptic

masking, and binocular summation, are expressed as,

$$binsum_T = \frac{stage1_{L_{TEST}}}{1 + \omega_D stage1_{R_{MASK}}} + \frac{stage1_{R_{TEST}}}{1 + \omega_D stage1_{L_{MASK}}}, \tag{4.3}$$

where ω_D is the weight of dichoptic masking, and *stage1* refers to equation 4.2, with subscripts indicating eye of origin and mask or test channel. The model is shown diagramatically in Figure 4.10B, with typical behaviour in Figure 4.10C. The best fit of this model for subject TAY, who shows the greatest half-binocular paradox, is shown in Figure 4.11B and E. Note that only the architecture of the model has changed - it has the same number of free parameters (3) as Model 1.

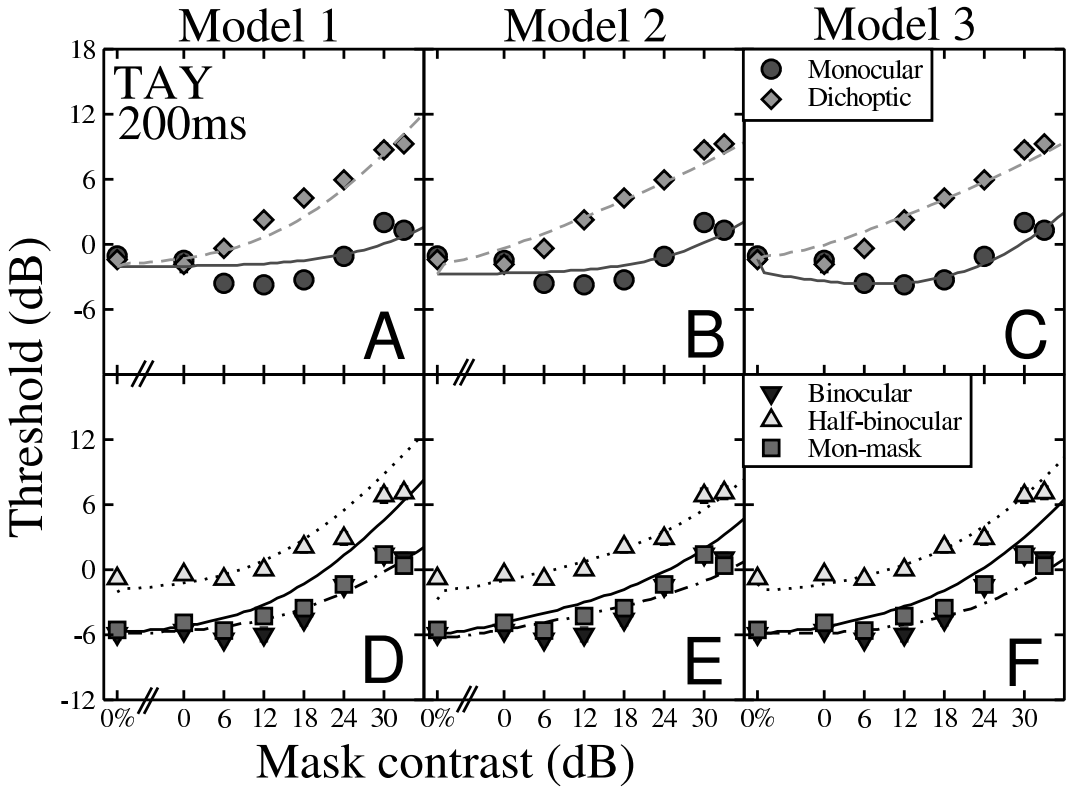

Figure 4.11: Best fits of Models 1 - 3 to the data of subject TAY. As previously, the models were fit to the monocular and dichoptic data (top row), and the parameters used to predict performance in the other conditions (lower rows). Parameters and error statistics of the fits were, for Model 2: ω_M=0.02, ω_D=0.51, k=0.007, RMSe=1.12dB, AIC=10.53, and for Model 3: ω_M=0.05, ω_D=0.73, k=0.02, α=9.68, RMSe=1.03dB, AIC=9.18. The fit of Model 1 is duplicated from Figures 4.8 and 4.9 for comparison.

Model 2 does a good job of describing half-binocular and binocular the results (Figure 4.11E). This is particularly impressive, as the model was fit only to the monocular and dichoptic functions, which means that the half binocular function is a prediction rather than a fit. For the other observers' data sets, the fits are comparable to those of Model 1, and in some cases are slightly superior (not shown).

It is also possible to model the facilitation observed in the monocular condition at 200ms duration. Numerous configurations were explored, based on models extant in the literature which place facilitation as a multiplicative term on the numerator of the gain control equation (Yu *et al.*, 2002; Meese and Holmes, 2007). Facilitation at stage one was never successful, however facilitation at stage two, after binocular summation, produced acceptable results. Stage two now becomes,

$$resp = \frac{({binsum_T}^p)(1 + \alpha \cdot binsum_M)}{Z + {binsum_T}^q}, \tag{4.4}$$

where *binsum* refers to the output of equation 4.3, in either the test T or mask M channels, and α is the weight of facilitation. The best fits of this model, which has four free parameters, and is termed Model 3, are shown in the right hand column of Figure 4.11 (panels C and F). The model does describe the facilitation observed for the 1cpd mask at 200ms, but leads to little or no improvement in the other conditions. The AIC score is lower for this model, indicating that the additional free parameter increases the descriptive capabilities of the model. For other subjects, where facilitation is not observed, α can be set to zero, and the model is identical to Model 2. The facilitation model is used again in Chapter 5.

4.5.3 Temporal components

Temporal dependencies can be introduced into the model in order to describe the results of Experiment 6. The monocular and dichoptic functions are not direct translations of the baseline detection thresholds, so three temporal functions should be required to represent baseline detection thresholds, and monocular and dichoptic masking functions. Since the data set shows strong evidence of the half binocular paradox, but no evidence of facilitation, Model 2 from the previous section is used (the absence of facilitation at high mask contrasts means that there is no way to constrain the value of α in Model 3).

Detection thresholds reduced as a function of stimulus duration, reaching asymptote at around 100ms (Legge, 1978; Luntinen *et al.*, 1995). Several mathematical functions were explored to represent this, including that of Luntinen *et al.* (1995). The simplest expression which describes the data adequately is an inverse exponential function,

$$f_E(t) = 1 - e^{-t/\tau}, \tag{4.5}$$

where t is stimulus duration in milliseconds, and τ is a free parameter. Monocular masking decreased with duration, and is described well by,

$$f_M(t) = t^{\eta}, \tag{4.6}$$

where t is again duration in milliseconds, and η is a free parameter, with a negative value for decreasing functions.

These functions are applied at stage one, in both the mask and test channels, to modulate the observer's sensitivity to the test stimulus. Stage one now becomes,

$$stage1_{L_{TEST}} = \frac{f_E(t){C_L}^m}{S + f_E(t)C_L + f_E(t)C_R + f_M(t)\omega_M X_L}, \tag{4.7}$$

where all terms are as for equations 4.2, 4.5 and 4.6. Binocular summation and dichoptic masking occur as previously (equation 4.3), followed by a standard version of stage two (without facilitation; equation 3.7).

The dichoptic masking process does not require a further temporal dependency, over and above what it is subjected to in the mask version of stage one, given by $f_E(t)$ in equation 4.5. Allowing an additional free parameter does not greatly improve the quality of the fit (Baker *et al.*, 2007c; Baker and Meese, 2006a,b). The reasons for this are explored in greater depth in Chapter 5.

The free parameters for this version of the model are τ and η which determine temporal dependencies, k, which governs overall sensitivity, and ω_M and ω_D, which control the weight of masking. Separate estimates of ω_M and ω_D were made for each mask type, with the other parameters kept constant for both masks. There were seven free parameters to fit five functions (monocular and dichoptic for two masks, and also detection thresholds). The half binocular data were not included in the fitting procedure or in calculating the reported error statistic, and are predictions from the other fits.

Best fits of the model are shown in Figure 4.12, with parameters and RMS errors given in Table 4.2. As for the previous experiment, the model was fit to detection thresholds, and the monocular and dichoptic functions. It is clear that the model describes the data well, producing small RMS errors ($<$ 1dB). The curve describing performance in the half binocular condition is a zero-free-parameter prediction, and follows the data closely in five out of six cases. The Two Stage model is clearly able to accommodate the results of this experiment, with minimal alterations.

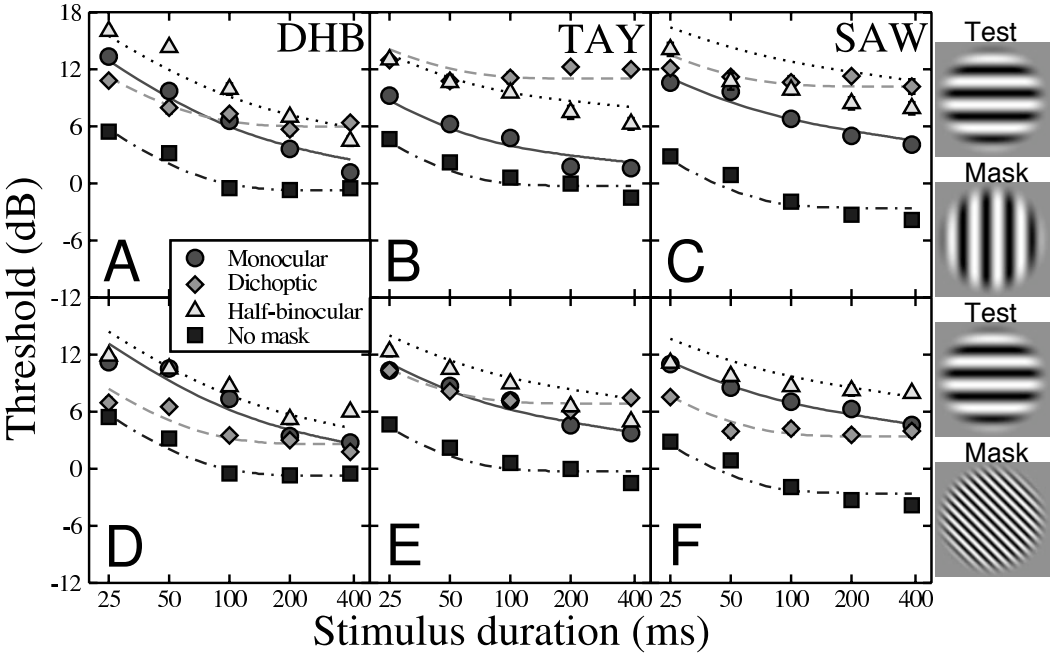

Figure 4.12: Temporal model fits to the data of Experiment 6. Layout is the same as for Figure 4.5, with curves giving the best fit of the model. Parameters were estimated as described in the text, and are given in Table 4.2.

Subject	RMS error (dB)	ω_{M1}	ω_{D1}	ω_{M3}	ω_{D3}	τ	η	k
DHB	0.84	1.03	0.25	1.09	0.12	38.68	-0.69	0.03
TAY	0.75	0.19	0.44	0.36	0.26	27.94	-0.46	0.04
SAW	0.68	0.36	0.59	0.37	0.24	30.74	-0.33	0.01

Table 4.2: Best fitting parameters for the temporal model. Fixed model parameters are as given in the caption to Table 4.1. Fits are shown in Figure 4.12. Weight parameters (ω) are indexed by their subscripts, where M is monocular, D is dichoptic, and numbers refer to the spatial frequency of the mask (either 1 or 3cpd).

Model weights and tuning

On a final note, a comparison of the weight parameters across mask condition is informative with respect to the spatial tuning of monocular and dichoptic masking. For monocular masking, the weights are similar within each subject across the two mask types (see Table 4.2, columns 3 and 5). This suggests that monocular cross-channel masking has broad spatial tuning, as changing the spatial properties of the mask has little effect. Dichoptic masking, on the other hand, is much weaker with the 3cpd oblique mask, and the model weights are approximately halved in value compared to those for the 1cpd orthogonal mask. This suggests that dichoptic cross-channel masking is more tightly tuned. This issue is considered in greater depth in Chapter 5.

4.6 Extending the Twin Summation Model

In Chapter 3, the Two Stage and Twin Summation models were found to be very similar. There was little to discriminate between them, although the Twin Summation model tended to be marginally superior. It has just been shown that the Two Stage model can be easily modified to incorporate a range of empirical data on cross-channel masking. However, the Twin Summation model has a very different architecture, which is much less flexible.

The first obvious place for cross-channel suppression to impact is at the gain control stage. However, this is after binocular summation in this model, and so any such modification would require monocular and dichoptic masking to be of the same magnitude, due to the logical argument presented in Section 4.5.1. This possibility has already been ruled out logically, and was confirmed computationally by implementing the model (not shown).

The alternative arrangement is for cross-channel masking to occur before binocular summation, in either the excitatory or suppressive channels, or both. This would require the introduction of some sort divisive stage, most likely of the form,

$$signal = \frac{C_L}{1 + \omega_M X_L + \omega_D X_R}, \tag{4.8}$$

where terms are as for previous equations. This configuration does a fair job of describing the empirical data from Experiments 5 and 6 (not shown), although it should be noted that with the addition of an early divisive stage, the model begins to strongly resemble the Two Stage model.

The main shortcoming of such a model is that it cannot produce the half binocular paradox. The most plausible account of this phenomenon is that it is caused by interocular

suppression within the mask channel, before it impacts the test mechanism. However, in the Twin Summation model, within-channel dichoptic masking is caused by the difference in exponents in the excitatory and suppressive pathways, and does not occur until the gain control stage, which is after binocular summation. This means that, even with some sort of feedback process, the dichoptic signal alone could not be modulated in such a way as to cause the paradox.

The Twin Summation model can therefore be rejected, on the grounds that it cannot encompass cross-channel effects without substantial modification, and that even with such modifications it cannot account for all of the empirical data. The model will not be considered further in this thesis.

4.7 Conclusions

This chapter is the first in depth exploration of cross-channel masking using different ocular configurations of mask and test. The key findings are that monocular and dichoptic masking have different magnitudes, temporal dependencies and spatial tuning, strongly suggesting that they arise from different processes. Both processes impact prior to binocular summation, and can be represented well by models of contrast vision which incorporate an early divisive gain control. The half binocular paradox indicates that additional suppressive processes may occur in dichoptic masking, or that dichoptic cross-channel suppression may occur after within-channel suppression.

Despite performing well, the models developed above are essentially functional descriptions of the data. They were not designed with neuroanatomical considerations in mind, and are not intended to be biologically plausible. Biological factors are addressed in Chapter 5, which begins with a review of relevant findings from the physiology literature.

Chapter 5

Further cross-channel masking experiments, influenced by neurophysiology

5.1 Physiology of cross-channel suppression

In neurophysiology, cross-channel suppression occurs when a cell's response to an optimal target grating is reduced by the presence of a mask distant in spatial frequency and/or orientation (Morrone *et al.*, 1982; DeAngelis, Robson, Ohzawa and Freeman, 1992; Bonds, 1989). This process could serve to normalize the cell's response (Heeger, 1992), perhaps preventing it from saturating at low contrasts (Meese and Holmes, 2007). However, there is no definitive account of the purpose of cross-channel suppression, and other interpretations have been proposed such as sharpening the cell's tuning properties (Blakemore and Tobin, 1972; Bonds, 1989). Initially, it was supposed that suppression was caused by inhibition between populations of cortical neurones preferentially tuned to different stimulus attributes (Morrone *et al.*, 1982; Heeger, 1992). However, a number of recent studies on cross-channel suppression in cat have cast doubt on this hypothesis, as will now be discussed.

When stimuli are presented monocularly, and mask and test are spatially and temporally superimposed, the tuning of the resulting suppression is not consistent with a cortical origin. Monocular cross-channel masking occurs at flicker rates too high for cortical neurones to respond (Freeman *et al.*, 2002; Li *et al.*, 2005; Sengpiel and Vorobyov, 2005; Sengpiel, Jirmann, Vorobyov and Eysel, 2006). Adaptation, which affects cortical, but not precortical cells (Movshon and Lennie, 1979; Ohzawa *et al.*, 1982, 1985), does not reduce monocular masking when the subject is first adapted to the mask (Freeman *et al.*, 2002). Many research groups have concluded that monocular masking must occur prior to V1, most likely in the lateral geniculate nucleus (LGN). Initial work proposed depression at the thalamocortical synapse as the mechanism by which cross-channel suppression occurs (Freeman *et al.*, 2002;

Carandini, Heeger and Senn, 2002). However, more recent studies have found that the time-course of masking is inconsistent with synaptic depression, and suggest that masking arises from nonlinear properties of LGN cells (Li *et al.*, 2006; Priebe and Ferster, 2006; Bonin, Mante and Carandini, 2004; Bonin *et al.*, 2005).

Dichoptic masking, on the other hand, is reduced after adapting to the mask (Sengpiel and Vorobyov, 2005; Li *et al.*, 2005). It is also reduced by introducing the GABA antagonist bicuculine into V1 (Sengpiel and Vorobyov, 2005; Sengpiel *et al.*, 2006). Intracortical inhibition is known to be mediated by GABA-ergic synapses. Furthermore, dichoptic suppression displays temporal frequency tuning similar to that of cortical neurones (Li *et al.*, 2005; Sengpiel and Vorobyov, 2005; Sengpiel *et al.*, 2006). These findings are consistent with a cortical origin for dichoptic suppression, which is further supported by fMRI evidence in human (Buchert, Greenlee, Rutschmann, Kraemer, Luo and Hennig, 2002).

A similar conclusion of two distinct suppressive mechanisms has also been reached in macaque for centre and surround masking (Webb, Dhruv, Solomon, Tailby and Lennie, 2005; Smith, Bair and Movshon, 2006). Masks falling outside the classical receptive field (CRF) of a neurone have similar properties to dichoptic masks, in that they are susceptible to adaptation, and have sharp spatiotemporal tuning. Central (overlayed) masks were also explored in these studies, and behave like monocularly presented masks described above. Psychophysically, there are also similarities; Petrov *et al.* (2005) report that surround suppression occurs prior to overlay suppression, and again shows the narrower tuning. However, there is no direct evidence to suggest that surround masking and dichoptic masking are mediated by the same neural process.

The findings from physiology have implications for the psychophysical study and modelling of cross-channel masking[1]. Crucially, the key findings from animal studies are yet to be demonstrated in humans. Some of these are explored in the following experiments. Experiment 7 was designed to address the order of monocular and dichoptic masking. Experiment 8 measures interactions between masking and adaptation. The models from Chapter 4 are then revisited in light of these experiments. Experiment 9 reports the spatial and temporal tuning of orthogonal masking, and Experiment 10 investigates the spatiotemporal dynamics of monocular and dichoptic processes.

[1] One consideration is the suitability of the term cross-channel masking. If masking arises due to saturation within circular filters, it is not occurring between channels. However, for consistency with the literature, the term is retained throughout the thesis.

5.2 Experiment 7 - cross-channel twin mask experiment

In Chapter 4, monocular and dichoptic cross-channel suppression were shown be distinct processes, both occurring prior to binocular summation. Whilst this is consistent with the findings from single-cell physiology discussed above, there is no psychophysical evidence for the serial ordering of the two suppressive sites implied by the LGN/V1 division. Such evidence is potentially important for further development of computational models, as well as for relating the findings from physiology to behavioural measures in human.

One experiment which could reveal the relative ordering of monocular and dichoptic suppression is as follows: a dichoptic pedestal masking (dipper) function is measured both with and without a cross-channel mask in the same eye as the pedestal. If monocular suppression occurs before dichoptic suppression, the cross-channel mask should reduce the effectiveness of the pedestal and shift the dipper function to the right. However, if dichoptic suppression occurs before monocular suppression, the cross-channel mask should not affect the pedestal before it suppresses the test, and the function should not shift to the right. In both schemes, the cross-channel component would also be expected to raise detection thresholds (by suppressing the test dichoptically), shifting the dipper function upwards.

Care must be taken in selecting stimuli for such an experiment. In particular, the ideal cross-channel mask is one which produces strong monocular masking (so that the cross-channel mask will powerfully suppress the pedestal) but weak dichoptic masking (so that the test will be directly affected only minimally). Data from Chapter 4 (see Figure 4.5) suggest that a 3cpd oblique mask at 50ms duration provides the required levels of cross-channel masking.

5.2.1 Methods

Apparatus and Stimuli

Dichoptic masking functions for a 1cpd horizontal grating were measured using a mirror stereoscope. As well as the standard dichoptic pedestal (0%; -10 to 30dB), there was an additional masking component presented dichoptically to the test. This additional mask was obliquely oriented ($-45°$), with a spatial frequency of 3cpd and a contrast of either 0% or 18% (25dB). Stimuli were presented using a ViSaGe.

Procedure

Experimental procedures were as described previously. Stimuli were presented for 50ms, with a 500ms ISI. Conditions were blocked by pedestal contrast and mask type. One full repetition was carried out at all 10 pedestal levels for each oblique mask contrast (0% or 25dB) before moving on to the next. Within each condition, a pair of staircases tracked responses when the test was in either the left or right eye. The experiment was repeated four times.

For this experiment, there were two mask components and one test component. This produces a problem for the frame interleaving technique (described in section 2.3), as interleaving three stimuli at 120Hz would produce noticeable flicker (since 40Hz per stimulus is below the critical flicker frequency of foveal vision). For this reason, composite (plaid) stimuli were used for the two mask components, in all conditions of the experiment (including those in which the cross channel component had a contrast of 0%).

Observers

DHB, TAY and SAW, the observers from the cross-channel experiments of Chapter 4, all participated in this experiment.

5.2.2 Results

The results of this cross-channel twin mask experiment are shown in Figure 5.1. The pedestal only condition (open symbols) is equivalent to the dichoptic condition of Experiment 1 (Chapter 3) and produces comparable results. This is reassuring, as the present experiment used a mirror stereoscope, whereas the previous experiment used shutter goggles.

The results for the cross-channel mask condition (filled symbols) are not particularly conclusive. There is a small rightward shift for all subjects, evident at the 10dB pedestal contrast for DHB, and the higher pedestal contrasts for TAY and SAW. This shift is most likely caused by the oblique mask reducing the effectiveness of the pedestal mask. However, the oblique mask has also raised detection thresholds in its own right; all subjects show threshold elevation at the lowest five pedestal contrasts. Although the effects are small, the data can be used for quantitative comparison of different model architectures, which are considered in the following section.

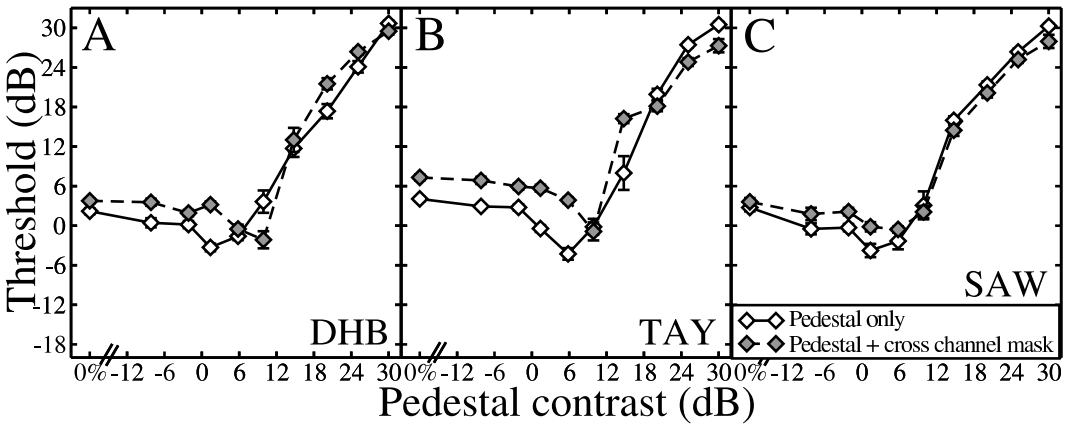

Figure 5.1: Results of the cross-channel twin mask experiment. Open symbols show a standard dichoptic pedestal masking function, replicating the findings in Chapter 3. Filled symbols are for the same pedestal, but with an additional 3cpd oblique mask at 25dB contrast, presented to the mask eye (dichoptic to the test).

5.2.3 Models

In Chapter 4 it was shown that monocular and dichoptic masking must impact prior to binocular summation. This restricts the set of possible model architectures. However, the order in which monocular and dichoptic masking impact is not known from previous psychophysical experiments. This is explored here, by considering several possible configurations, and comparing their ability to describe the data of the above experiment.

In the Two Stage model there are three likely locations where cross-channel masking might impact: before, after, and at stage one. There are also two masking processes, monocular and dichoptic, yielding $3^2 = 9$ possible models to be considered. Table 5.1 shows the locations at which masking impacts for each model (note that the numbering of models does not reflect the models discussed in previous chapters). These are divided into three sub-groups, the first of which is consistent with findings from the neurophysiology literature discussed in section 5.1. The groups are:

- models in which monocular masking occurs before dichoptic masking,
- models in which monocular masking occurs after dichoptic masking,
- models in which both kinds of masking impact at the same location.

The equations describing each model all have the same basic format. The following series of expressions sets out the response for a single ocular channel (here, left) of the test mechanism,

$$r_{iL} = \frac{C_{iL}}{1}, \qquad r'_{iL} = \frac{r^m_{iL}}{S + r_{iL} + r_{iR}}, \qquad r''_{iL} = \frac{r'_{iL}}{1}, \tag{5.1}$$

Model number	Before S1	At S1	After S1
1	M	D	-
2	-	M	D
3	M	-	D
4	D	M	-
5	-	D	M
6	D	-	M
7	MD	-	-
8	-	MD	-
9	-	-	MD

Table 5.1: Configurations of monocular (M) and dichoptic (D) cross-channel masking. S1 is stage one of the Two Stage model. Suppression is always divisive, as outlined in equations 5.1, 5.3 and 5.4.

where the i subscript denotes the test channel and stimuli, C is stimulus contrast, L and R indicate left and right eyes, and m and S are model parameters (see Chapter 3). For these equations, where cross-channel masking is absent, the first and third expression are clearly superfluous, and the model is formally equivalent to the Two Stage model when the following equation is used to compute output,

$$resp = \frac{(r''_{iL} + r''_{iR})^p}{Z + (r''_{iL} + r''_{iR})^q}, \tag{5.2}$$

where p, q and Z are model parameters, discussed further in Chapter 3. This equation is common to all models considered in this section[2]. Figure 5.2 shows how the four expressions of equations 5.1 and 5.2 map onto the standard two stage model. Note that although there are now four 'stages' to the model, only two of them are fully fledged contrast gain control stages which feature self suppression, exponentiation and a saturation constant which can differ from unity. For this reason, the model is still referred to as the two stage model.

Cross-channel suppressive terms are then included on the denominators of equation series 5.1, according to the order shown in Table 5.1. Thus, Model 1 (for the left channel), in which monocular masking occurs before stage one, and dichoptic masking occurs at stage one, is given by,

$$r_{iL} = \frac{C_{iL}}{1 + \omega_M C_{jL}}, \qquad r'_{iL} = \frac{r^m_{iL}}{S + r_{iL} + r_{iR} + \omega_D r_{jR}}, \qquad r''_{iL} = \frac{r'_{iL}}{1}, \tag{5.3}$$

where j indicates the mask channel and stimuli. As a further example, Model 9, in which both masking processes occur after stage one, is,

[2]The facilitatory processes discussed in Chapter 4 are ignored here as they cannot be constrained by the available data.

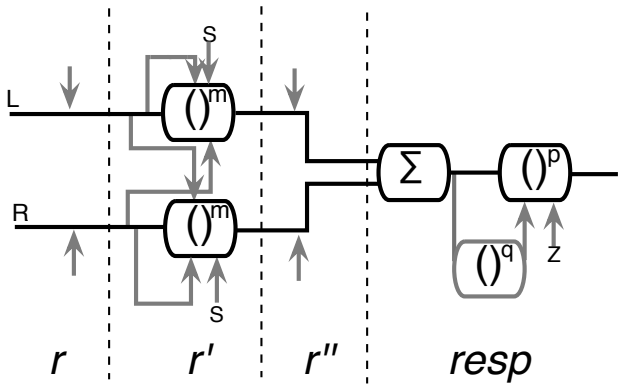

Figure 5.2: Model diagram showing where cross-channel suppression can impact on the two stage model. The *r, r'* and *r"* terms are the three parts of equation 5.1, and the *resp* section is the output of equation 5.2. Unlabelled grey arrows are locations at which cross-channel suppressive signals can impact.

$$r_{iL} = \frac{C_{iL}}{1}, \qquad r'_{iL} = \frac{r_{iL}^m}{S + r_{iL} + r_{iR}}, \qquad r''_{iL} = \frac{r'_{iL}}{1 + \omega_M r'_{jL} + \omega_D r'_{jR}}. \tag{5.4}$$

Note that the cross-channel terms which occur after the first equation are fed by the r'_j terms, not the input mask contrast C_j. This means that the mask terms have been through the equivalent first and/or second equations in the jth channel. This point will be returned to later in the chapter.

In the absence of dichoptic cross-channel suppression ($\omega_D = 0$), the nine models simplify to three configurations (shown diagramatically in the upper portion of Figure 5.3), which produce subtly different rightward shifts of the dipper function (Figure 5.3, lower panels). When monocular cross-channel suppression occurs before stage one (Fig. 5.3A), the cross-channel mask shifts the dip to the right, but keeps the handles parallel. When suppression occurs at stage one (Fig. 5.3B), the handles converge at high contrasts. Finally, when suppression occurs after stage one (Fig. 5.3C), the handles diverge at high contrasts. Although the predicted functions do appear quite similar, the data might allow the most appropriate configuration to be determined.

In order to best discriminate between the models, as many parameters as possible should be given fixed values. In the absence of cross-channel components, all models are identical to the standard Two Stage model. For this reason, the parameters *p, q, m* and *Z* are fixed at values from Meese *et al.* (2006) for all subjects. Furthermore, for each data set, an initial fit to the pedestal only condition (0% mask contrast) is used to estimate k and S, so that these can be kept constant also. Thus, to fit the twin mask data, the only free parameters are the two suppressive weights, ω_M and ω_D. As all models have the same fixed and free

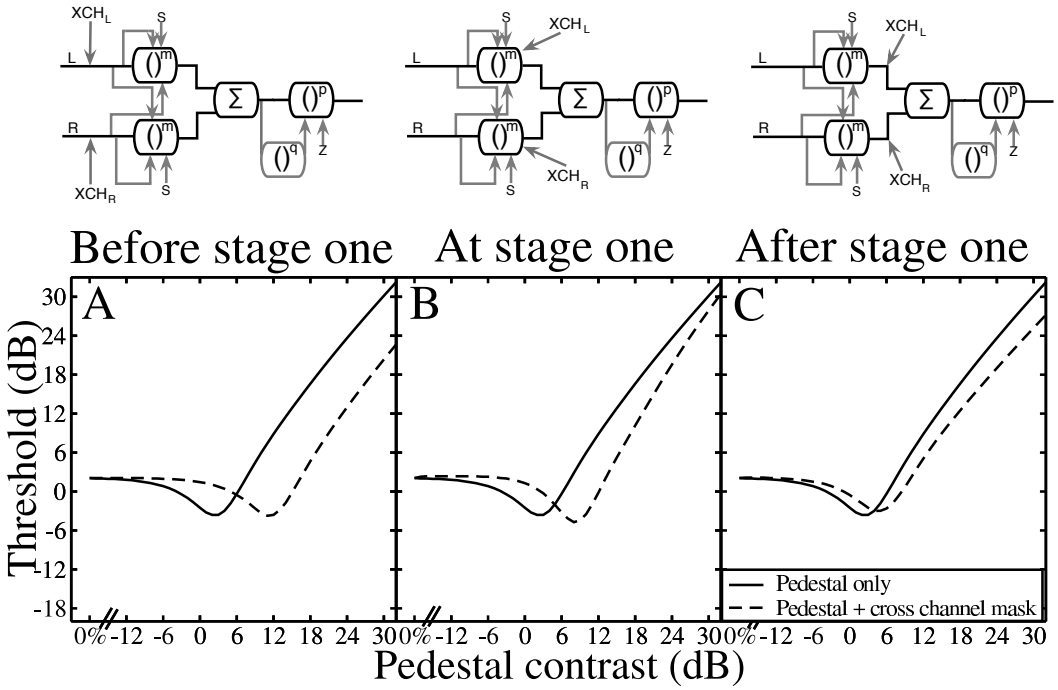

Figure 5.3: Model predictions for three locations of monocular cross-channel masking. The cross-channel mask has a contrast of 25dB, with $\omega_M = 0.1$. Masking can occur either before (A), at (B) or after (C) stage one. The solid lines are standard dichoptic pedestal masking functions, and the dashed lines are for pedestal masking with an additional cross-channel mask dichoptic to the test. The impact of the cross-channel mask on the test is not considered here ($\omega_D = 0$), as this can only raise the function upwards, regardless of where dichoptic suppression impacts.

parameters, the only performance determining factors will be the locations where masking impacts, as laid out in Table 5.1.

RMS errors for all nine models (and for the fit of the standard Two Stage model to the pedestal data only, referred to as Model 0) are shown graphically in Figure 5.4A. Symbols give the best fits for each subject, and also to the data averaged across all three subjects (stars). The fit to the pedestal data only (Model 0) is good for DHB and the averaged data, and is acceptable for subjects TAY and SAW. For the full data set, all nine models are similar, particularly for the averaged data and subject SAW, for whom the range in RMS values is <0.5dB. In general, most fits are fair, (RMS errors <2dB), although there are some very poor fits for subject TAY.

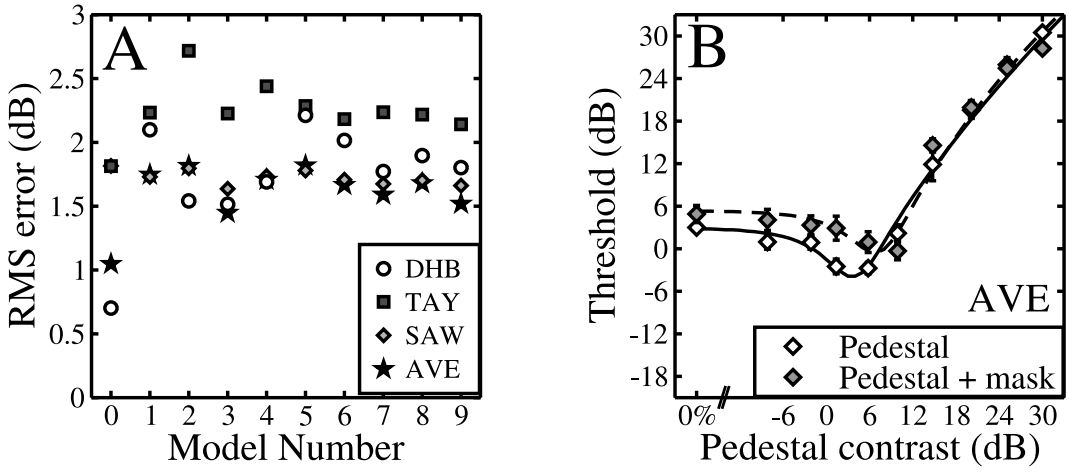

Figure 5.4: RMS errors for nine model configurations, fit to the data of Experiment 7, along with an example fit. A) Symbols correspond to individual subject fits, or a fit to the averaged data (across all three observers; stars). Model 0 is the best fit to the pedestal data alone (open symbols in Figure 5.1), which is used to constrain parameters for the fits to the pedestal + mask condition. Other models are variants of the Two Stage model, with monocular and dichoptic suppression impacting at different locations, as shown in Table 5.1. B) The best fit of Model 3 to the averaged data across 3 subjects.

Overall, the best model is Model 3, in which monocular masking occurs before stage one, and dichoptic masking occurs after it. This model produces the smallest RMS error for the averaged data set, and subjects DHB and SAW. The fit is also good for TAY, however Model 9 produces a marginally superior fit for this subject (by around 0.08dB). Figure 5.4B shows the best fit of Model 3 to the data averaged across 3 subjects. This is the fit which produced the lowest RMS error across all data sets (1.45dB).

5.2.4 Discussion

A masking experiment was carried out using either one or two mask components. The first was a dichoptic pedestal, and the second was a dichoptic cross-channel mask. The

cross-channel mask had a small effect on the dichoptic (pedestal) dipper function, shifting it upwards and to the right. The ability to account for this behaviour was tested for nine variants of the Two Stage model, with monocular and dichoptic suppression occurring in different locations. By a very small margin, the best fitting model was one in which monocular suppression occurs before stage one, and dichoptic suppression occurs after stage one (Model 3).

The best model is consistent with the arrangement of monocular and dichoptic effects inferred by physiologists (see section 5.1). It is also consistent with the most highly developed model in Chapter 4, because dichoptic suppression occurs late. This means that the mask signal has already passed through stage one in its own mechanism, and been subjected to within-channel dichoptic masking, and is therefore able to produce the half-binocular paradox in the manner discussed in section 4.5.2.

Although this experiment provides the first psychophysical evidence concerning the order of monocular and dichoptic cross-channel suppression, it is not especially conclusive. It is unfortunate that the effects measured proved to be so small, as a larger effect would have allowed for more powerful discrimination between models. Larger effects could be achieved using a higher contrast cross-channel mask, although this would reduce the number of available pedestal contrasts, given the limited dynamic range of the equipment. Instead, an adaptation paradigm is used in the following experiment to further localise the two masking processes in human.

5.3 Experiment 8 - adaptation of cross-channel masking

Until recently, the dominant account of cross-channel masking was that of inhibition between cortical mechanisms. Since cortical neurones are prone to adaptation (Movshon and Lennie, 1979; Ohzawa *et al.*, 1982, 1985), one might expect the level of masking observed psychophysically to be reduced after adapting to the mask. However, previous psychophysical studies have failed to find this effect using binocular presentation (Foley and Chen, 1997; Holmes, 2003), which at the time was somewhat puzzling.

The single-cell findings discussed above (see section 5.1) suggest an explanation for this. Since monocular cross-channel suppression is apparently pre-cortical, it should not be prone to adaptation. Dichoptic suppression, on the other hand, is most likely cortical, and so masking should be reduced after adapting to the mask. Indeed, this result has been reported for cat physiology (Sengpiel and Vorobyov, 2005; Li *et al.*, 2005). In the following experiment, cross-channel masking is measured psychophysically after first adapting to the mask, for both monocular and dichoptic presentation.

5.3.1 Methods

Apparatus and Stimuli

The adaptor was a vertical sinusoidal grating with a spatial frequency of 1cpd, which filled the circular stereoscope aperture (9° diameter) for one eye only, and had a contrast of either 50% (adapt) or 0% (no adapt). The other eye viewed mean luminance throughout the adaptation period. The adaptor was shifted in phase every 100ms by a random amount, $180° \pm n$, where $0 < n < 90°$. This was to reduce local luminance adaptation, and also to prevent subjects from making eye movements to track the grating (Holmes, 2003).

Stimuli were static (100ms), and as described for previous experiments (1cpd, horizontal test, vertical mask), presented using a mirror stereoscope and a ViSaGe.

Procedure

On a given day, only one eye was adapted, either to the 50% or the 0% adaptor. Thresholds were measured for the conditions in which the mask was presented to the adapted eye (one monocular and one dichoptic condition), at all mask contrasts (0%; 6 to 30dB, depending on subject). Experimental sessions were blocked by mask contrast. Each block started with a 2 minute adaptation period before trials began. Each trial (see Figure 5.5) consisted of a 6 second adaptation 'top up' period, during which the adaptor was again displayed, followed by a 500ms blank interval before stimulus presentation. Stimuli were presented for 100ms with a 500ms ISI. Subjects were instructed to complete the block (around 10-15 minutes) without pausing, so as to avoid reducing the level of adaptation. Data were collapsed across eye of presentation, and across all four repetitions of the experiment, prior to analysis.

Observers

Three subjects, DHB, TAY and SAW participated in this experiment. SAW was measured at two mask contrast levels, TAY at four, and DHB at six. Each subject completed four full repetitions of the experiment (for each eye) at each mask contrast tested.

5.3.2 Results

In Figure 5.6, the upper panels show the monocular results, and the lower panels show the dichoptic results. In each panel, open symbols indicate performance with a 0% adaptor, and filled symbols with a 50% adaptor. It is clear that there is no substantial reduction in the level of masking for monocular presentation. Indeed, for subject DHB there appears to

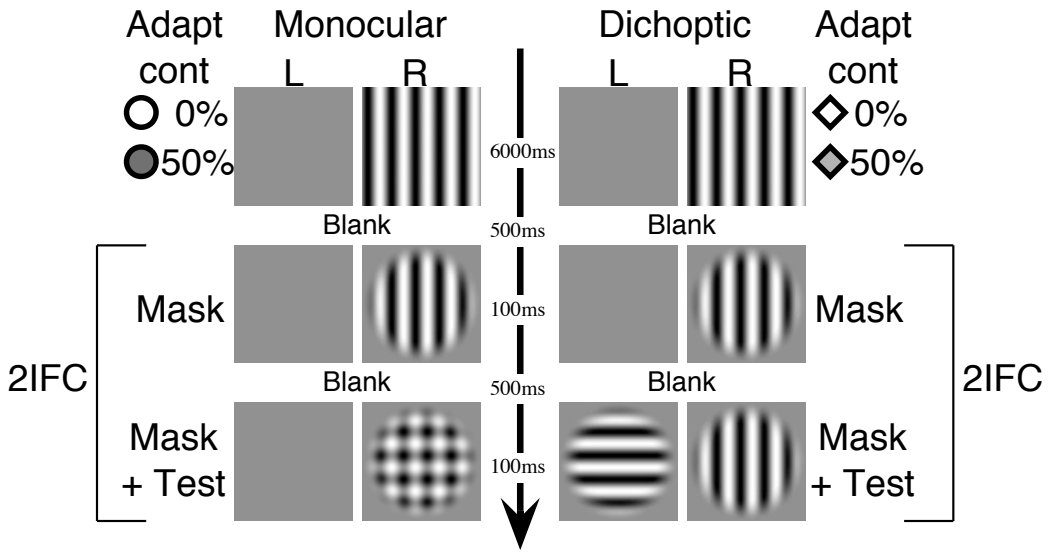

Figure 5.5: Temporal sequence for each trial of the adaptation experiment. A trial began with 6 seconds of 'top-up' adaptation, followed by stimulus presentation, which could be either monocular (left) or dichoptic (right). The mask was always presented to the adapted eye, which was constant on a given day. The order of monocular/dichoptic presentation, and mask/mask+test, was randomised.

be an increased level of masking after exposure to the high contrast adaptor. This strongly suggests that the process which causes monocular masking is not weakened by adaptation.

The dichoptic data show a different pattern. For all subjects, the level of dichoptic masking is greatly reduced after adaptation, by as much as 5dB (a factor of 1.78). This is a consistent result across all mask contrasts above 0%, for all subjects. Interestingly, subjects DHB and SAW also show a small reduction at detection threshold (0% mask contrast) for the test after adapting to an orthogonal grating in the other eye. This is considered further in the Discussion.

5.3.3 Discussion

The effect of adaptation on contrast masking was explored for orthogonal masks, presented monocularly and dichoptically. For three subjects, the level of monocular masking was not reduced after adaptation, but the level of dichoptic masking was substantially weakened. These findings support the conclusions from the previous chapter that monocular and dichoptic masking are caused by separate processes. The dichoptic effect is consistent with a cortical origin, as it is well established that cortical neurones show reduced activation after adaptation (Movshon and Lennie, 1979; Ohzawa *et al.*, 1982, 1985). The absence of this effect for monocular presentation argues against a cortical locus, consistent with the hypothesis that monocular masking arises in the LGN (Freeman *et al.*, 2002; Li *et al.*, 2005, 2006; Priebe and Ferster, 2006). This result provides important evidence that the conclusions

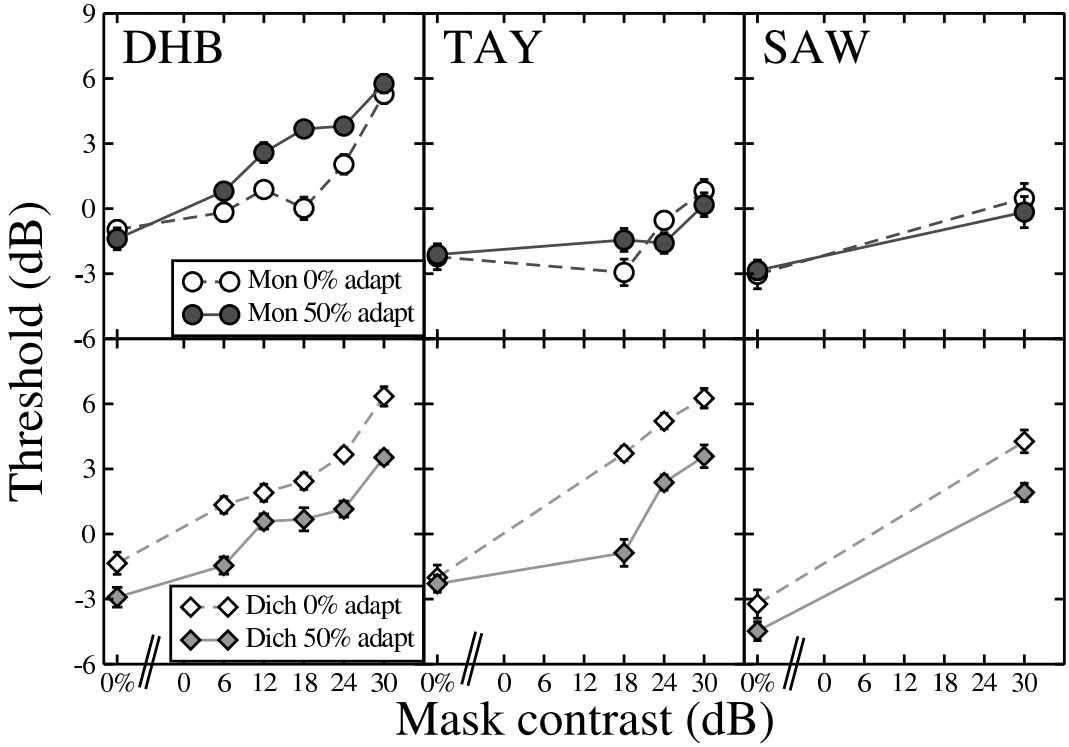

Figure 5.6: Results of the adaptation experiment. Upper panels show performance in the monocular condition, with either 0% (open) or 50% (filled) adapting contrast. The adaptor does not reduce the level of masking substantially, and even raises it for one subject (DHB). Lower panels show performance in the dichoptic condition. Here, adapting to the mask does reduce the level of masking for all three subjects, and at all mask contrast levels.

drawn from single-cell experiments in cat (and discussed in Section 5.1) might also apply to the human visual system.

One unexpected result is that for two subjects, the baseline detection threshold (no mask) was reduced after adapting to an orthogonal grating in the other eye. Given that the adaptor presumably lies far outside the orientation bandwidth of the mechanism most sensitive to the test, this is quite surprising. One possible explanation is that a standing level of (dichoptic) suppression has been reduced by the adaptor, as suggested by Morrone *et al.* (1982). Although cortical cells produce little spontaneous firing when not stimulated (Hubel, 1959), it is possible there might be enough activity to cause weak dichoptic suppression, even in the absence of a mask stimulus. If this is the case, the main conclusion is therefore supported by the observation that a similar effect is not found monocularly.

5.3.4 Modelling

The best model from Experiment 7 was used to fit the data. For the test mechanism, the model is described by the following equations:

$$r_{iL} = \frac{C_{iL}}{1 + \omega_M C_{jL}}, \qquad r_{iR} = \frac{C_{iR}}{1 + \omega_M C_{jR}}, \tag{5.5}$$

$$r'_{iL} = \frac{{r_{iL}}^m}{S + r_{iL} + r_{iR}}, \qquad r'_{iR} = \frac{{r_{iR}}^m}{S + r_{iR} + r_{iL}}, \tag{5.6}$$

$$r''_{iL} = \frac{r'_{iL}}{1 + \omega_D r'_{jR}}, \qquad r''_{iR} = \frac{r'_{iR}}{1 + \omega_D r'_{jL}}, \tag{5.7}$$

$$resp = \frac{(r''_{iL} + r''_{iR})^p}{Z + (r''_{iL} + r''_{iR})^q}, \tag{5.8}$$

where ω_M and ω_D are monocular and dichoptic weights, C is contrast of stimuli in either the left (L) or right (R) eye, for either optimal (test) stimuli (i) or non-optimal (mask) stimuli (j), and all other symbols (m, S, p, q, Z) are model parameters as described in Chapter 3.

Thus, the first equation (5.5) represents cross-channel monocular masking, the second equation (5.6) is the familiar gain control stage from the standard Two Stage model (equation 3.6), the third equation (5.7) produces cross-channel dichoptic masking, and the final equation (5.8) gives both binocular summation and the second stage of the Two Stage model. Note that an equivalent set of equations govern activation in the mask channel, and are solved to give values for r'_j in equation 5.7.

As adaptation is often considered as an inhibitory (Dealy and Tolhurst, 1974) or gain control effect (Wilson, 2003), it is assumed that adapting a mechanism increases the saturation constant (Heeger, 1992), S, in the adapted eye (equation 5.6 for the mask channel). This lowers activation in the adapted channel, reducing dichoptic masking (equation 5.7), which occurs after adaptation, but not monocular masking (equation 5.5), which occurs before adaptation. Of course, adaptation is undoubtedly a more complex process, or series of processes (Langley, 2002), than can be exhaustively modelled by changing a single parameter[3]. For example, Foley and Chen (1997) rejected several adaptation models in which only the saturation constant was varied, and found that it was necessary to also vary a further (sensitivity) parameter to describe their data (see also Meese and Holmes, 2002). Adding an additional sensitivity parameter to the present model produced little or no improvement in the fit, and a model in which only the sensitivity parameter (and not the saturation constant) was varied produced poorer fits. This parsimonious, single parameter account of adaptation is therefore sufficient to describe the results here.

In fitting the data, four free parameters were allowed to vary: k, S_A, ω_M and ω_D. The term S_A denotes the saturation constant after adaptation, which was re-estimated for the mask mechanism in the adapted eye. Other terms are as previously, and the remaining model parameters (*p, q, Z, S, m*) were fixed at the values from Meese *et al.* (2006).

Unlike the modelling in previous sections, here a sequential fitting procedure was used. An estimate of k was obtained first, by minimising the error between the model predictions and the threshold data. Then estimates of ω_M and ω_D were calculated by fitting the model to the data from the appropriate 0% adapt contrast functions (open symbols). Finally, an estimate of S_A was calculated by fitting to the adapted data.

This sequential procedure is possible because, unlike the models discussed previously, the parameters are largely independent of each other. Sequential fitting has several advantages over the standard simultaneous multidimensional parameter estimation procedures. Firstly it is much faster, with fits taking seconds rather than minutes. Furthermore, it is less prone to problems with local minima, common with multidimensional algorithms. This means that the optimal parameters are generally found first time, so less time is spent performing multiple fits. Finally, the order of fitting constrains the model to always fit well at certain key points, such as baseline detection thresholds.

The best model fits are shown in Figure 5.7, with parameters and RMS errors given in Table 5.2. In all cases, S_A is much larger than S (S=0.985), indicating that adaptation increased the saturation constant in the adapted mechanism. The fits are very good, partic-

[3]Changing the saturation constant results in a rightward shift of the contrast response function. However, Albrecht, Farrar and Hamilton (1984) found both lateral and vertical (compressive) shifts of striate contrast response function after adaptation. Thus, modelling adaptation as a change in S is a simplification, but one which is adequate for describing the present data.

Subject	RMS error (dB)	k	ω_M	ω_D	S_A
DHB	1.14	0.017	0.038	0.315	15.26
TAY	0.56	0.011	0.012	0.406	19.08
SAW	0.48	0.005	0.021	0.393	15.92

Table 5.2: Parameters and error statistics for the adaptation model. The parameters were estimated by a sequential fitting procedure, as described in the text. Fits are shown in Figure 5.7.

ularly in the dichoptic condition, which suggests this is an appropriate method of modelling adaptation. In particular, the shape of the masking function after adaptation is captured well, becoming steeper and more curved than with a 0% adaptor. RMS errors are excellent, particularly for TAY and SAW, although for SAW this is hardly surprising given the small number of data points included in the fit.

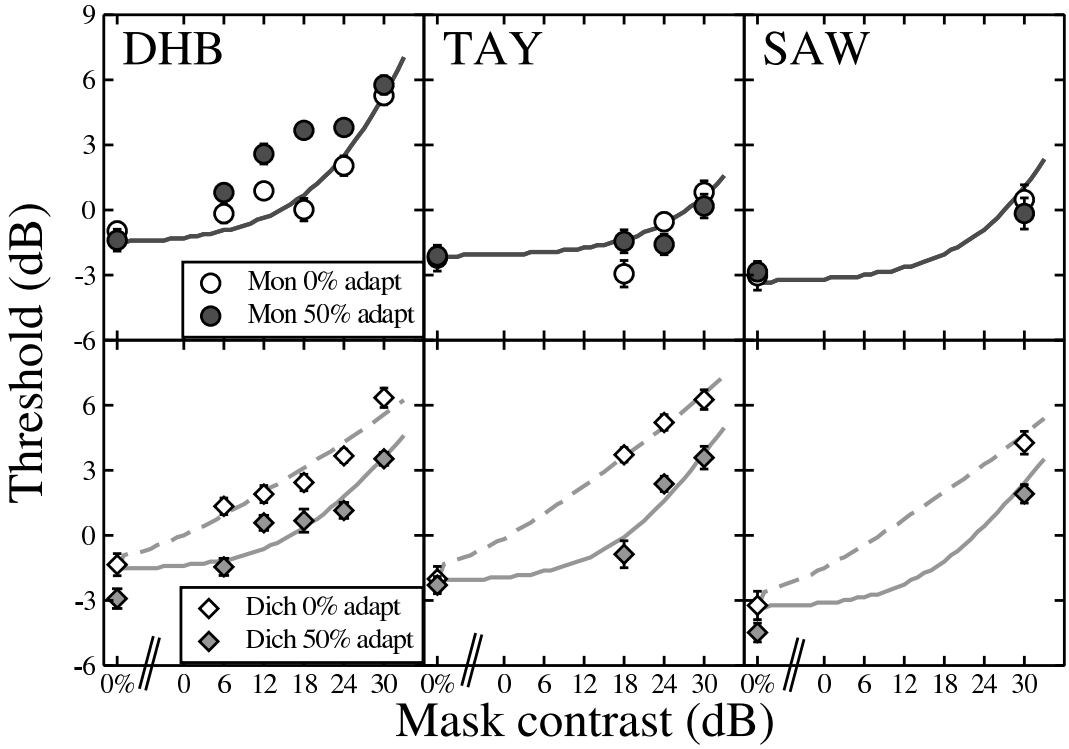

Figure 5.7: Model fits to adaptation data. Data are replicated from Figure 5.6, and curves are the best fits of the adaptation model, as described in the text. Parameters and error statistics are given in Table 5.2.

The main shortcoming of the model is in the monocular condition. For DHB, the increase in masking after adaptation is not captured. Also, for TAY, there is a small amount of facilitation observed prior to adaptation, consistent with the results in Chapter 4, which is not captured by this implementation of the model. These two features may be related. If cross-channel facilitation is indeed a late acting process, as the modelling in Chapter 4 has suggested, then it would be affected by adaptation, in the direction observed for subject DHB (note that although DHB shows no evidence of facilitation *per se*, it might still be

present but be obscured by the strong masking, as suggested by Meese and Holmes (2007)). This assertion is not tested here, as there is too little data to accurately constrain the weight of facilitation in the model. However it seems the most likely explanation of the monocular results.

Several alternative model configurations can be rejected by these data. Any models in which adaptation occurs exclusively at a late stage (for example, after binocular summation) would not affect the suppressive mechanisms which produce monocular and dichoptic masking of differing magnitudes (see Chapter 4). Furthermore, any model in which adaptation occurred very early on (prior to any masking effects) would predict an adaptation aftereffect for the monocular condition, which is clearly not observed. Thus, the site of adaptation important for this experiment can be placed after the point at which monocular cross-channel suppression occurs (believed to be the LGN) but before the point at which dichoptic suppression impacts (most likely primary visual cortex).

Finally, this model arrangement is also consistent with a widely studied adaptation phenomenon - interocular transfer. The interocular transfer of adaptation occurs when sensitivity to a target is reduced by first adapting to that target presented at high contrast in the other eye (Blakemore and Campbell, 1969; Bjorklund and Magnussen, 1981; Blake, Overton and Lema-Stern, 1981a). Typically, the aftereffect is about 60% of that for presentation in the adapted eye. If adaptation is assumed to occur not only at the monocular stage one in our model, but also again at later binocular stages (i.e. stage two), then this effect could easily be explained (see Blake *et al.*, 1981a, for detailed discussion of this scheme). Adaptation is strongest in the adapted eye, because both the monocular and binocular sites are adapted. When the non-adapted eye is tested, the monocular stage one has not been affected, but the later binocular stage is desensitised. Thus, adaptation will still occur, but will be weaker than in the adapted eye. A quantitative assessment of this scheme would require further detailed experimentation.

5.4 Biologically plausible models

The models developed so far have been largely motivated by psychophysical data, and a desire to keep them as simple as possible. However, given what is known about the architecture of the visual system in general, and about the neural substrates of cross-channel masking, other model architectures might be more appropriate from a biologically plausible standpoint. There are two obvious areas in which this applies. Firstly, given the conclusions of Li *et al.* (2006) and Priebe and Ferster (2006) that monocular cross-channel suppression occurs in the LGN, the processes underlying monocular cross-channel masking effects

are reconsidered in the following section. Secondly, the implications of including feedback processes in the model are explored.

5.4.1 Monocular masking

For the reasons discussed in Section 5.1, monocular cross-channel masking is believed to be a pre-cortical phenomenon. Initially, this was ascribed to depression at the thalamocortical synapse (Freeman *et al.*, 2002; Carandini *et al.*, 2002). However, two recent papers have suggested that this is unlikely, as the time-course of cross-channel suppression is not consistent with the time-course of synaptic depression (Li *et al.*, 2006; Priebe and Ferster, 2006). Instead, both studies present evidence that suppression occurs due to nonlinearities in LGN neurones, specifically saturation and thresholding.

Priebe and Ferster (2006) also present a simple yet convincing feedforward model which implements this architecture. Briefly, a number of LGN neurones are simulated, with orientationally isotropic receptive fields. These respond to both the mask and the test, and tile the available stimulus area. The LGN neurones feed into higher units, resembling cortical simple cells, the orientation tuning of which is determined by the spatial locations of the LGN neurones sampled.

If the LGN neurones are linear, then the entire system behaves linearly, and masking is not observed (the principle of superposition). However, if a saturating nonlinearity is introduced to the LGN cells, such as the Naka-Rushton function (Naka and Rushton, 1966), then the increase in the input to the cortical units caused by the test will be smaller with additional (mask) components present than without. This is because each LGN neurone will have been moved out of the accelerating (steep) part of its response function, and into the saturating (shallow) region. An increase in test contrast will produce a smaller increase in output than if the mask were absent, and suppression (or masking) will have occurred. The model produces convincing simulations of the contrast response of cortical neurones to optimal (test) stimuli in the presence of non-optimal (mask) stimuli. Because the suppression is within-channel, it is similar to the process suggested by Legge and Foley (1980) in their influential early model of masking (see section 1.2.3).

It is not clear how best to convert a neuronal model such as that of Priebe and Ferster (2006) into a system-level model which is useful in describing psychophysical data. A recent attempt by Chirimuuta and Tolhurst (2005b) to produce dipper functions from a population of model neurones was partly successful, but at the expense of reconciling the distribution of neuronal sensitivities between the model and empirical data. An alternative possibility is to produce a fully fledged image processing model, into which could be fed the images used as experimental stimuli (i.e. Watson and Solomon, 1997). Such a model might confer much

predictive power, particularly when considering complex masks with many components, and is discussed further in Chapter 8.

The approach taken here is to use a simple mathematical operator which can account for the breadth of psychophysical data. Divisive suppression has all the required properties, so it is retained for the remainder of the thesis. However, it should be stressed that this is not the only possible process[4] and is a generalisation rather than a strictly accurate model. It is also clear that, in order to be faithful to what is currently known about the neurophysiology, monocular suppression should occur before dichoptic suppression, as it does in the preferred model from Experiment 7.

5.4.2 Feedback models

Up to this point, only feedforward model architectures have been considered. This is a common approach to psychophysical modelling (Ding and Sperling, 2006a,b), as it avoids the complexities necessary to implement the dynamic neural models (i.e. Wilson, 1999, 2003) required for feedback. However, committing to feedforward architectures also presents problems, as has occurred with the current preferred model.

In Chapter 4, an extra stage of dichoptic masking was added to the Two Stage model, in order to account for the half-binocular paradox. This paradox is the empirical finding that orthogonal dichoptic masking can be reduced by adding a further orthogonal mask in the same eye as the test. A second stage of dichoptic suppression (after stage one) predicts this phenomenon because the extra mask suppresses the dichoptic mask 'within-channel' at stage one, *before* it has affected the test. The dichoptic signal is reduced, and subsequently produces weaker masking.

However, there is no evidence from either psychophysics or neurophysiology to suggest that there are two distinct mechanisms of dichoptic suppression, one within and one between channels. Indeed, such an arrangement is quite unlikely, as the cross-channel mechanism would have to be tuned to all channels except that of the test mechanism, and would thus require a weighting function with a notch in fourier space. A more plausible account is one in which dichoptic masking impacts at a single site (stage one) within the context of a feedback model. Such an arrangement should also produce the half-binocular paradox, as the cross-channel dichoptic signal will be affected by the extra mask at the same point at which it suppresses the test.

To demonstrate this, the existing model was recast as a simple dynamic feedback model. Examples equations are given in terms of the left test mechanism, but apply also to the

[4]Subtractive models have been used successfully in both vision (Langley, 2002; Manahilov, Gordon, Calvert and Simpson, 2007) and hearing (Moore and Glasberg, 1996).

right eye, and to mechanisms tuned to mask stimuli. The input to the feedback portion of the model is the input contrast, inhibited by any masking components in the same eye (see equation 5.5 above);

$$r_{iL(0)} = \frac{C_{iL}}{1 + \omega_M C_{jL}}, \tag{5.9}$$

where C denotes contrast in the test (i) or mask (j) channels, and ω_M determines the weight of monocular masking. In order to represent a continuous system using a discrete computer representation, the state of the system is sampled iteratively. The subscript number (0 in the above equation) is the iteration number. The feedback portion of the model encompasses stage one,

$$r'_{iL(t)} = \frac{(r_{iL(t-1)} + r_{iL(0)})^m}{S + (r'_{iL(t-1)} + r'_{iL(0)}) + (r'_{iR(t-1)} + r'_{iR(0)}) + \omega_D r'_{jR(t-1)}}, \tag{5.10}$$

where t is the iteration number, and r'_{jR} is the mechanism sensitive to mask stimuli in the right eye. The model is then completed by binocular summation, and the second stage of gain control, as described previously in equation 5.2. A diagram of the model is shown in Figure 5.8.

Typically the model required a small number iterations (4 or 5) to reach asymptotic performance. An example of this is shown in Figure 5.9 for dichoptic masking using a 30dB cross-channel mask. It is clear that activation in both the test and mask channels has stabilised after five iterations. To avoid making any arbitrary assumptions about termination criteria, the model was run for ten iterations for all simulations.

With suitable parameters, the model predicts the half-binocular paradox for cross-channel masking (Figure 5.10A). Furthermore, performance for within-channel stimuli is not affected. Figure 5.10B shows monocular, binocular and dichoptic contrast discrimination functions which exhibit all of the main features of the models and data presented in Chapter 3. Thus, it can be demonstrated that a feedback model predicts all of the experimental phenomena associated with previous feed-forward models. There is also an architectural similarity to several models of binocular rivalry, most notably that of Lehky (1988), which also features an early stage of interocular feedback suppression (discussed further in Chapter 8).

The success of the feedback model raises the question of how to proceed with modelling in the remainder of the thesis. The feedback model has strong biological plausibility, as it is well known that feedback connections are abundant in visual cortex. However, it is computationally time consuming, and less straightforward to manipulate than previously

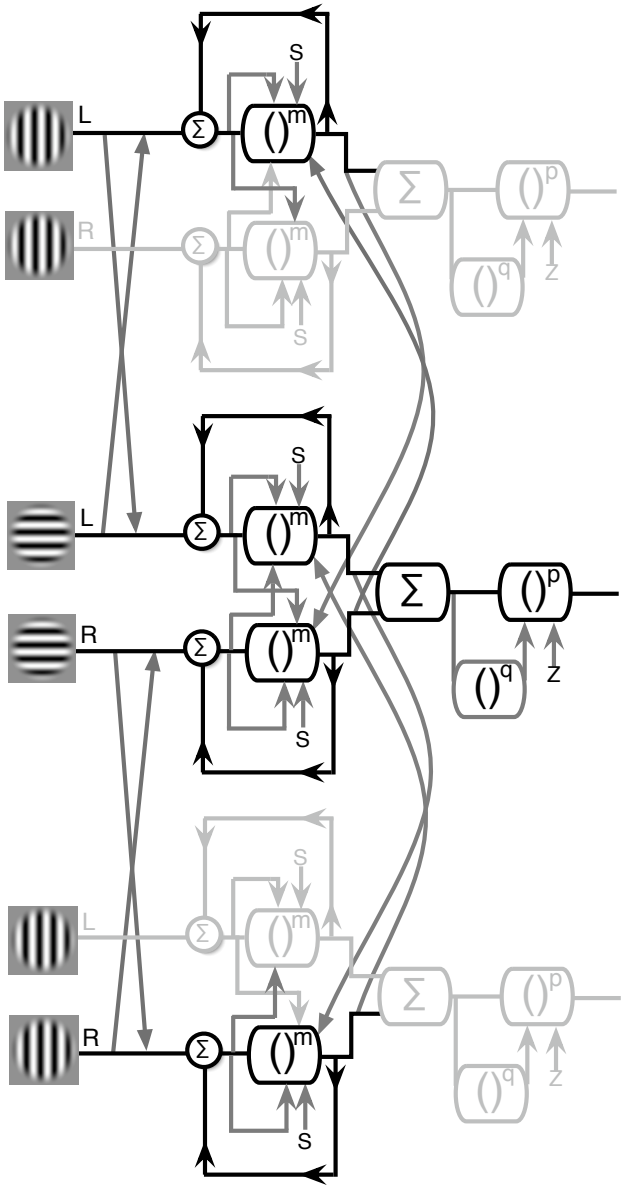

Figure 5.8: Diagram of the feedback model. The central system is the mechanism sensitive to the horizontal test stimulus. The flanking systems are mechanisms sensitive to orthogonal masks, and are shown twice for ease of exposition. Shaded out regions of the masking mechanisms are not required for understanding the feedback processes, but are included for completeness. Feedback occurs within and between channel. As for previous model diagrams, dark grey arrows indicate divisive suppression, Σ indicates summation of inputs, and brackets raised to a power indicate exponentiation.

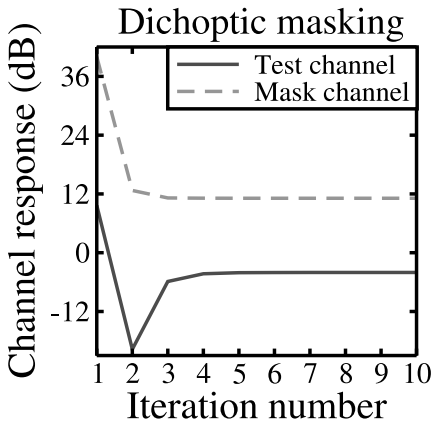

Figure 5.9: Channel activation in the feedback model. The model was stimulated with a test and a 30dB dichoptic cross-channel mask. Activation (the output of equation 5.10) is shown over 10 iterations in both the test and mask channel, and is converted to dB for clarity.

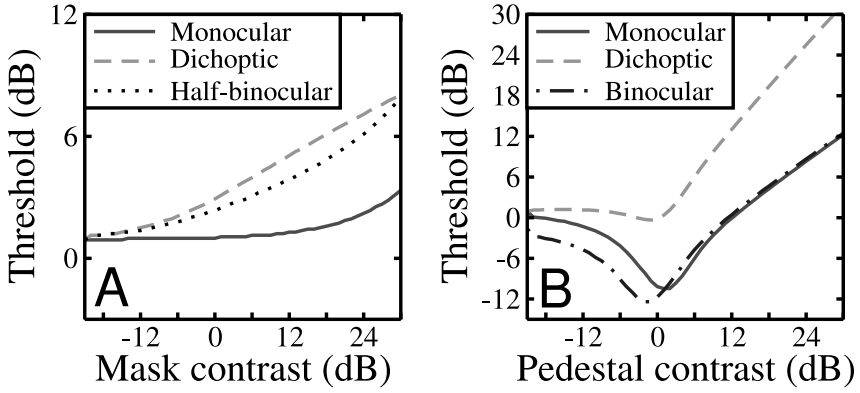

Figure 5.10: Behaviour of the feedback model. The left panel shows performance with cross-channel masks, and predicts the half-binocular paradox when monocular masking is weak. The right panel shows that performance for within-channel masks is unaffected. The parameters from Meese *et al.* (2006) were used, although S was increased slightly (from 0.98 to 1.5) to give more appropriate dipper placement. For the cross-channel case, $\omega_M = 0.01$ and $\omega_D = 0.9$, consistent with parameters from previous versions of the model.

developed feedforward models. A compromise is to continue using the feedforward model developed in section 5.2.3 (in which cross-channel dichoptic suppression impacts after it has been through stage one), but to bear in mind that this model is a simplification of what might well be a feedback system.

5.5 Experiment 9 - spatiotemporal tuning of orthogonal masking

Up to this point, several experiments have shown that cross-channel masking comprises two separate processes, linked to monocular and dichoptic masking. It is important to characterise these processes in terms of their spatial and temporal frequency tuning, which might well be different given the distinct neural substrates. The data in Chapter 4 offer a hint that monocular cross-channel masking might be more broadly tuned for spatial frequency than dichoptic masking. Monocular thresholds for a 1cpd test are elevated approximately equally by an orthogonal 1cpd mask and an oblique 3cpd mask. Dichoptic presentation, on the other hand, produces weaker masking for the 3cpd mask than for the 1cpd mask. Single-cell results from Li *et al.* (2005) suggest that a similar situation might be true of temporal frequency tuning. Experiment 9 explores the spatial and temporal frequency tuning of orthogonal masking for both monocular and dichoptic presentation.

5.5.1 Methods

Apparatus and Stimuli

Stimuli were sinusoidal gratings, spatially limited by a raised cosine envelope. The test spatial frequency was always a horizontal 1cpd grating, counterphase modulated at 4Hz. Mask gratings were always vertical, and varied in either their spatial or temporal frequency (the untested dimension was the same as for the test). For spatial frequency, the spatial envelope remained constant, and the mask gratings were varied in logarithmic steps of spatial frequency between 0.25 and 4cpd. For temporal frequency, eight waveforms were used between 2 and 22.6Hz. The waveforms were sinusoidal, and windowed by a raised cosine envelope, with a central plateau of 250ms. Peak mask contrast was always 30dB, and the time-averaged contrast for each waveform was between 45.7 and 52.4% of the peak for all waveforms. Stimulus examples, and temporal waveforms, are shown in Figure 5.11.

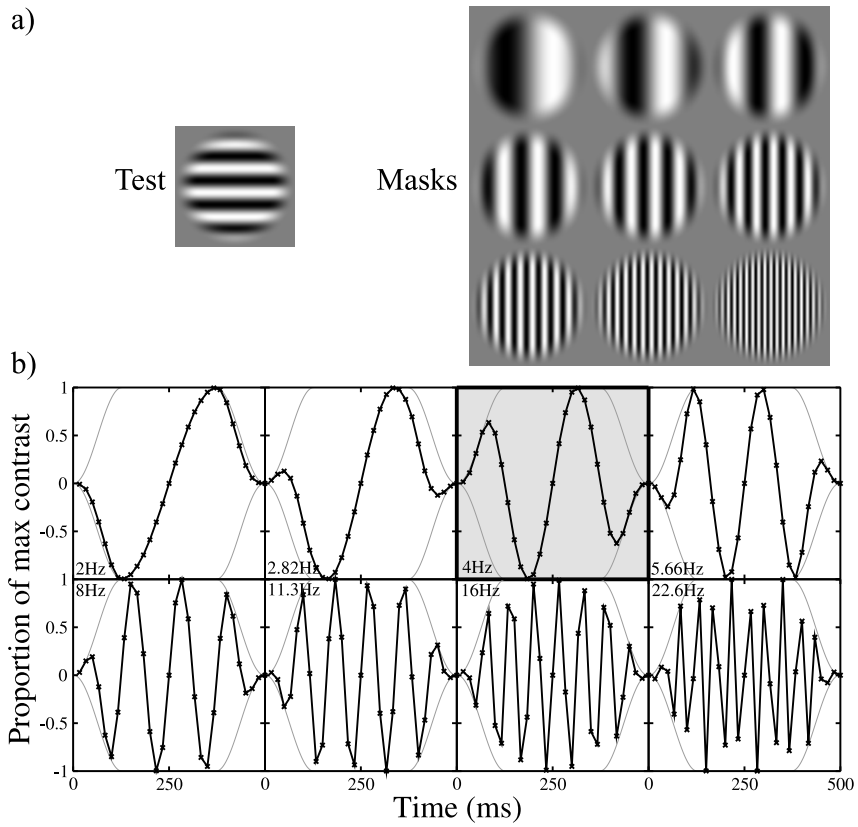

Figure 5.11: Stimuli and temporal waveforms for Experiment 9. a) shows examples of test and mask stimuli, and b) gives the temporal waveforms for all stimuli. Crosses in b) are the actual sample points for the monitor refresh rate used (120Hz). Faint grey lines show the raised cosine envelope. The highlighted panel indicates the test frequency (4Hz).

Procedure

Stimuli were presented using a standard 2AFC procedure, with a stimulus duration and ISI each of 500ms. Stimuli were blocked by mask frequency. Within each block, four staircases tracked performance for monocular and dichoptic presentation in both the left and right eye. All conditions in the spatial frequency experiment were always completed together, likewise for temporal frequency (i.e. the blocks were not interleaved between spatial and temporal functions). Both subjects also completed a baseline condition, in which detection of the test (horizontal, 1cpd, 4Hz) was recorded with no mask present. All conditions were repeated four times, and the results analysed as for previous experiments.

Observers

Subjects DHB and RJS completed the experiment wearing their normal optical correction.

5.5.2 Results

The spatial frequency tuning functions are shown in Figure 5.12. Threshold elevation is apparent at all mask frequencies, in both the monocular (filled circles) and dichoptic (open diamonds) conditions. The monocular functions are similar for both observers, and show very little dependency on mask frequency. Threshold elevation is around 3-6dB, depending on subject, over most of the range tested. There is a slight tendency for masking to decrease towards the highest spatial frequencies. Thus, monocular masking is spatially broadband.

Dichoptic masking is also broadly tuned, though less so than for monocular presentation. Above the test frequency, there is a steep decline of >3dB per octave. At lower mask frequencies, the function shows a shallower fall off for RJS, and is flat for DHB. Dichoptic masking is therefore more narrowly tuned for spatial frequency than monocular masking.

The temporal frequency tuning functions have a similar character to those for spatial frequency, and are shown in Figure 5.13. Again, masking is apparent at all frequencies, reaching almost 12dB (a factor of 4) in some conditions. The monocular functions show an increase with temporal frequency to a peak at a frequency higher than that of the test (11Hz for DHB and 16Hz for RJS). There is a reduction in masking as temporal frequency increases further. Again, monocular masking is very broadly tuned.

The dichoptic functions show a different pattern. For RJS, as temporal frequency increases, the amount of threshold elevation reduces, over a 6dB range across the 3.5 octaves tested. As for monocular masking, the sharpest decline is between the highest two frequencies. Subject DHB shows a less distinct tuning, maintaining around 6-9dB of threshold

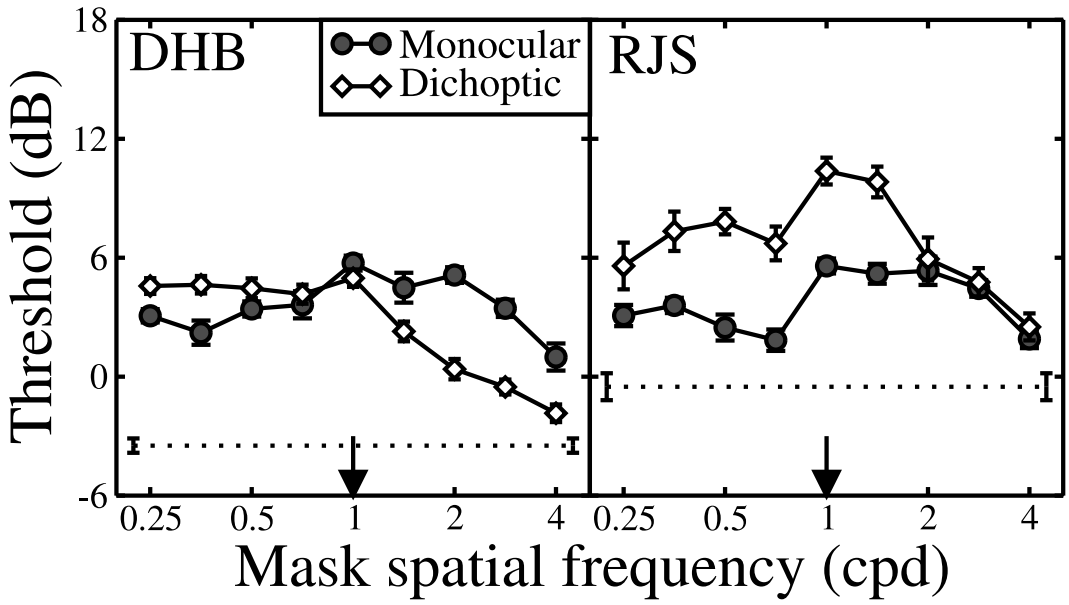

Figure 5.12: Spatial frequency tuning of orthogonal masking. Results are shown for two observers, DHB (left) and RJS (right). Data show monocular (filled circles) and dichoptic (open diamonds) detection thresholds for a horizontal test grating, in the presence of a vertical 30dB mask, over a range of mask spatial frequencies. The arrow indicates the test frequency (1cpd). All stimuli were counterphase modulated at 4Hz. Dotted lines give the baseline detection threshold (0% mask contrast), and error bars are the standard error of the probit fit.

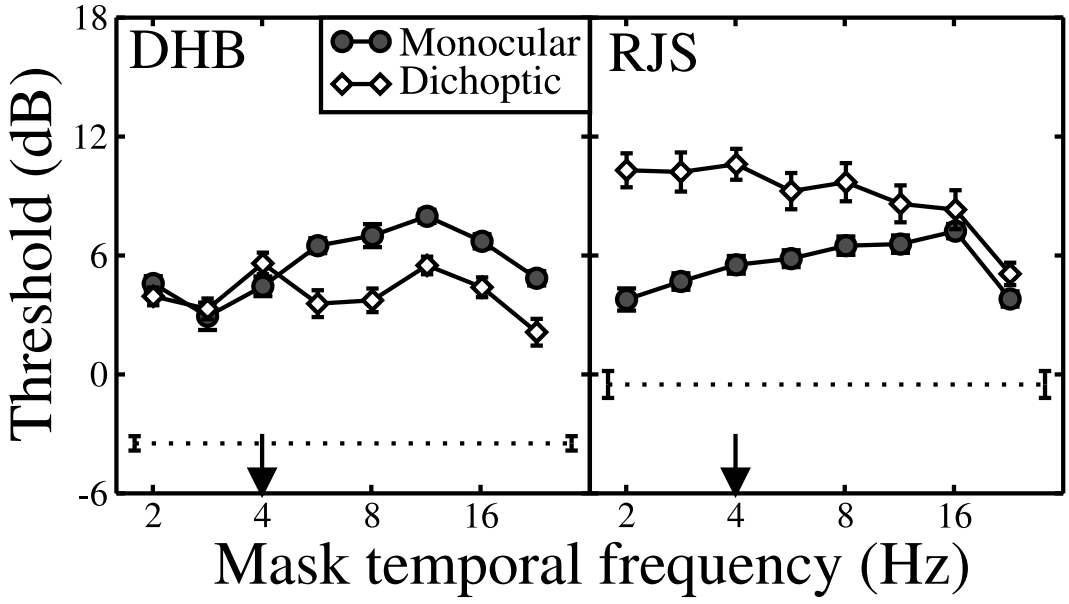

Figure 5.13: Temporal frequency tuning of orthogonal masking. Results are shown for two observers, DHB (left) and RJS (right). Data show monocular (filled circles) and dichoptic (open diamonds) detection thresholds for a horizontal test grating, in the presence of a vertical 30dB mask, over a range of mask temporal frequencies. The arrow indicates the test temporal frequency (4Hz). All stimuli had a spatial frequency of 1cpd. Dotted lines give the baseline detection threshold (0% mask contrast), and error bars are the standard error of the probit fit.

elevation across the frequencies tested. Dichoptic masking therefore also shows broad temporal tuning, but has different features from monocular masking.

5.5.3 Discussion

Two experiments were carried out to explore the spatial and temporal tuning of orthogonal masking, under monocular and dichoptic presentation. All functions showed broad tuning, with masking apparent over a 4 octave range of frequencies. Monocular functions had the broadest tuning, with dichoptic functions showing a high spatial frequency cut. Monocular functions peaked at mask temporal frequencies at least an octave above that of the test.

The broad monocular tuning is consistent with findings from single-cell physiology (Bonin *et al.*, 2005; Freeman *et al.*, 2002; Li *et al.*, 2005; Sengpiel and Vorobyov, 2005) and is also similar to psychophysical results using binocular presentation (Petrov *et al.*, 2005; Holmes, 2003; Meese, 2004; Meese, Holmes and Challinor, 2007a). Dichoptic masking has narrower tuning, again consistent with the findings from physiology (Li *et al.*, 2005; Sengpiel and Vorobyov, 2005; Sengpiel *et al.*, 2006), and the findings of Chapter 4. However, it is worth pointing out that the temporal frequencies at which dichoptic masking does not occur physiologically (Sengpiel and Vorobyov, 2005; Li *et al.*, 2005) are at or above the highest frequencies used here ($\geq$16Hz). It is therefore interesting that dichoptic masking is seen at the highest mask frequencies. There is some physiological evidence that some cortical neurones are suppressed by high temporal frequency stimuli (see Figure 1B of Li *et al.*, 2005). Although these cells are presumably greatly in the minority, they may be responsible for the small masking effects (<6dB) seen here at the highest temporal frequencies. Overall, it appears that monocular masking is broadly tuned, whilst dichoptic masking is narrower, but still quite broad, consistent with single cell studies (Li *et al.*, 2005; Sengpiel and Vorobyov, 2005; Sengpiel *et al.*, 2006). This tuning presumably represents the bandwidth of the underlying suppressive gain pool

An interesting side issue is to compare the magnitude of masking found here with the data of Chapter 4, in which static stimuli were used. For subject DHB, a 30dB orthogonal 1cpd mask presented for 200ms produced around 6dB of threshold elevation for dichoptic presentation, and even less for monocular presentation. However, with the temporally modulated stimuli used here, threshold elevation was around 9dB, when mask and test had equal spatiotemporal properties, and was even higher at other temporal frequencies. This is particularly surprising, given that the time-averaged contrast for temporally modulated stimuli is around 50% of the peak, so the total contrast energy of a temporally modulated mask is lower than that of a static mask. This comparison suggests that temporal modulation produces greater masking, perhaps by more strongly activating the magnocellular (transient)

pathway (Tolhurst, 1975b; Legge, 1978; Nachmias, 1967; Merigan and Eskin, 1986; Merigan, Byrne and Maunsell, 1991a; Merigan, Katz and Maunsell, 1991b; Lynch, Silveira, Perry and Merigan, 1992; Lennie and Movshon, 2005), which has been proposed as the source of monocular cross-channel suppression (Li *et al.*, 2006; Meese and Holmes, 2007).

5.6 Experiment 10 - scale dependence of cross-channel masking

Whereas Experiment 9 kept the test properties constant and varied only the mask, it is also possible to co-vary the spatiotemporal properties of the mask and test, to explore a different region of the stimulus space. Experiment 10 measures masking over a wide range of spatiotemporal scales. Using a similar paradigm, Meese and Holmes (2007) found that the weight of binocular cross-channel masking was directly proportional to the square root of the speed (TF/SF) of flickering stimuli. Here, it is assessed whether or not this rule holds for monocular and dichoptic presentation.

5.6.1 Methods

Apparatus and Stimuli

In order to allow for comparison across studies, this experiment used many of the same conditions as Meese and Holmes (2007). Stimuli were orthogonal patches of sinusoidal grating, windowed by a gaussian spatial envelope (a Gabor function). Spatial frequencies were 0.5, 1, 2 and 4cpd, and the gaussian envelope always had a full-width at half-height of 1.65 grating cycles.

Two temporal waveforms were used; 4 and 15Hz. The 4Hz envelope was a sine wave, multiplied by a raised cosine envelope with a total duration of 500ms and a 250ms central plateau. The 15Hz waveform was a biphasic pulse, with a total duration of 66.7ms (8 frames at 120Hz). Stimulus examples and temporal waveforms are shown in Figure 5.14.

All previous experiments in this thesis have used a single central fixation point. However, for small patches of high spatial frequency gratings, it is likely that such an arrangement could affect performance by interfering with the stimulus (Summers and Meese, 2007). A 'quad' arrangement of four points equidistant from the centre was used instead. The points were placed 3 cycles of the carrier grating away from the centre, so their configuration changed with stimulus frequency (Meese and Holmes, 2007).

A ViSaGe running in pseudo-14bit mode was used for one subject (DHB), and a VSG2/4

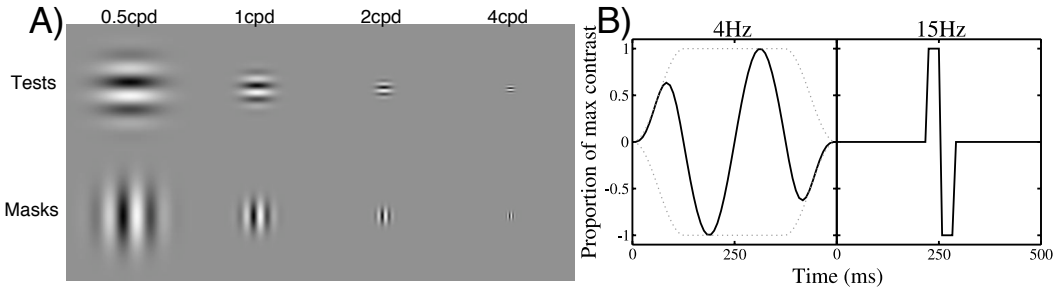

Figure 5.14: Stimuli and temporal waveforms used in Experiment 10. A) shows high contrast examples of test stimuli (top) and mask stimuli (bottom) at the four spatial frequencies used. B) shows the two temporal waveforms used (dark lines). The grey dotted line in the left panel is the raised cosine envelope.

running in pseudo-15bit mode was used for the other subjects. The same monitor and stereoscope arrangement was used for all subjects, and the stimuli and experimental software (Liberator) were also identical throughout.

Procedure

Stimuli were blocked by mask contrast, and spatio-temporal condition. Subjects completed one repetition of all mask contrasts and temporal frequencies at a given spatial frequency, before moving on to the next spatial frequency. Within each block, four staircases tracked monocular and dichoptic masking for both the left and right eyes, as reported previously. A single block took around either 4 minutes (15Hz) or 8 minutes (4Hz), owing to the different stimulus durations (see above). The ISI was 400ms for all conditions.

All observers repeated the experiment three times, and the results were merged across repetition and eye of presentation, before using probit analysis to estimate a threshold.

Observers

3 subjects completed this experiment, DHB (author) and two male undergraduate optometry students, WS and KP, both aged 20. The undergraduates were psychophysically naïve, and participated as part of their course requirements. All subjects were emmetropic.

5.6.2 Results

The raw data for all three observers is shown in Figure 5.15. The panels are arranged in a similar way to those of Meese and Holmes (2007). The highest speed condition (low SF, high TF) is in the lower left hand corner, and the lowest speed condition (high SF, low TF) is in the upper right hand corner. Within each panel, the horizontal dotted line indicates

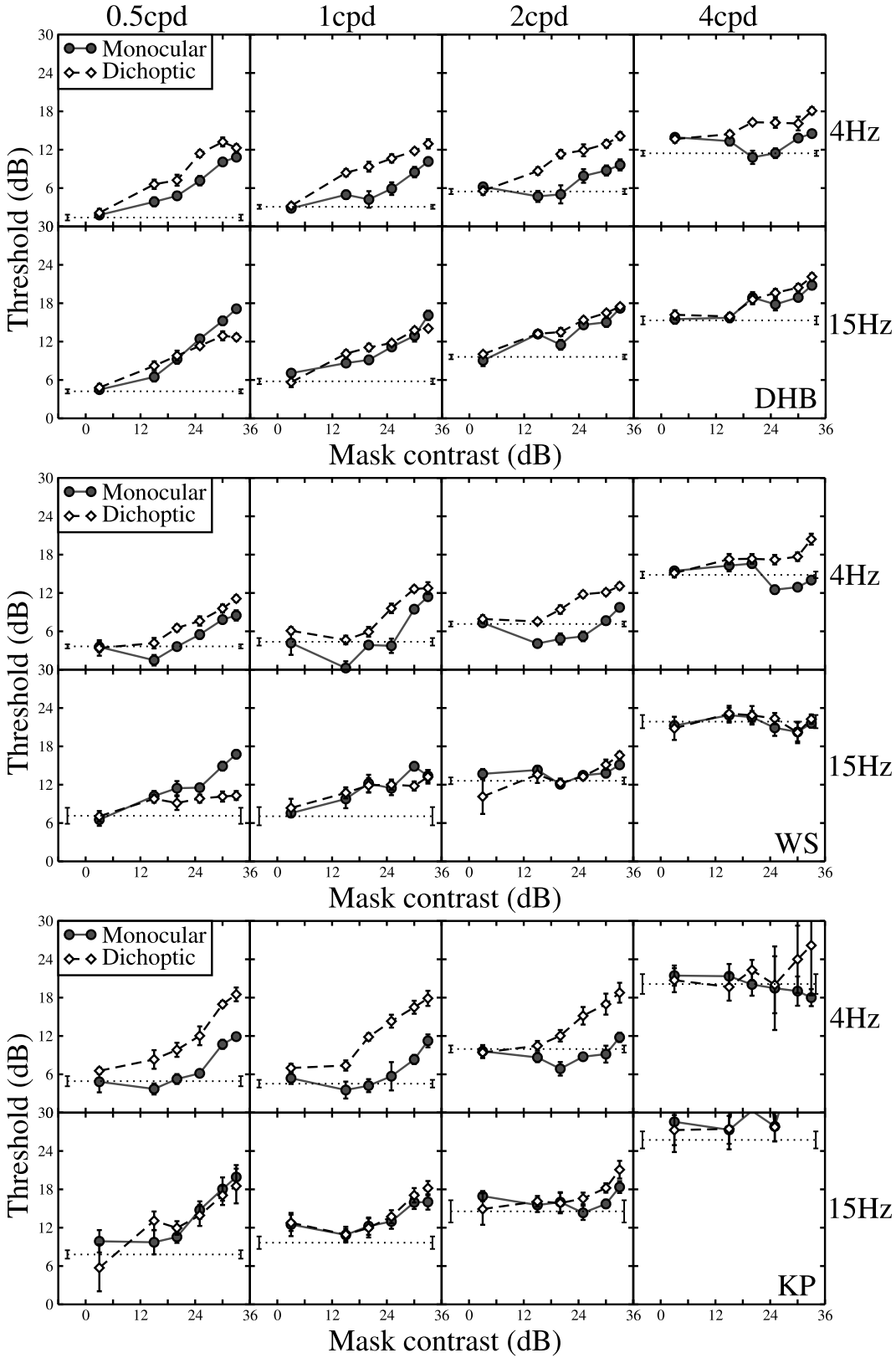

Figure 5.15: Monocular and dichoptic cross channel masking at different spatiotemporal frequencies. Spatial frequency increases by panel from left to right (see top plot), temporal frequency is 4Hz in the upper rows, and 15Hz in the lower rows of each plot. The three plots correspond to three observers, given by initials in the lower right panels. Error bars give the standard error of the probit fit, and horizontal dotted lies are baseline detection thresholds (mask contrast of 0%).

the baseline (0% mask contrast) detection threshold. This is lowest at spatial frequencies of 0.5 and 1cpd, and increases with spatial and temporal frequency.

Thresholds for monocular presentation (filled circles) generally increased with mask contrast, although a small amount of facilitation was apparent (discussed further below). Monocular masking appears strongest at high speeds (lower left corner), and is markedly reduced in the lower speed conditions (upper right corner). In general, monocular masking was weaker than dichoptic masking. However, for some conditions, this pattern was reversed, consistent with the results of Chapter 4 (i.e. 0.5cpd, 15Hz condition for WS).

The dichoptic masking functions (open diamonds) appear to show less of a dependency on speed. Masking is apparent for all spatiotemporal conditions, and although it appears weaker towards the upper right hand corner this might be due to the elevated baseline thresholds in this region. One straightforward method of assessing this is to normalize both axes to detection threshold. This allows for comparison across spatiotemporal condition, as the relative mask contrasts can be compared to the relative threshold elevation. Figure 5.16 shows the normalized functions, collapsed across spatiotemporal condition onto single panels for monocular and dichoptic masking for each observer[5].

Several interesting features are apparent from the normalized data in Figure 5.16. In the monocular condition, a spread of masking functions is seen, similar to that reported by Meese and Holmes (2007). This indicates that the monocular data are affected by the spatial and temporal dynamics of the stimuli. Furthermore, there is clear evidence of cross-channel facilitation for two subjects (WS and KP). This has a very different character from the within-channel facilitation discussed in Chapter 3. Facilitation from a pedestal is typically maximal when the pedestal contrast is around detection threshold. However, here the greatest facilitation is found at around 12dB above (four times) detection threshold. Furthermore, the facilitation seen here is always <6dB, whereas pedestal facilitation can be as great as 18dB (see Wichmann, 1999, and Figure 7.4), although is typically ∼9dB (see Figure 3.1).

No facilitation is evident in the dichoptic condition. However, unlike the monocular results, the dichoptic data appear to collapse onto a single function. This is particularly convincing for DHB and KP. Subject WS, whose data are quite noisy, shows less of a spread of masking functions than in his monocular data. This is good evidence that dichoptic masking is unaffected by the spatial and temporal frequency of the stimuli, and is therefore spatiotemporally scale invariant. This is explored further by computational modelling in the next section.

[5]For subject KP, the detection threshold at 4cpd, 15Hz, is extremely high. Thus, the thresholds for the masking function often fall outside the displayable contrast range of the equipment. This makes data for this condition unreliable, so they are omitted from all further analyses

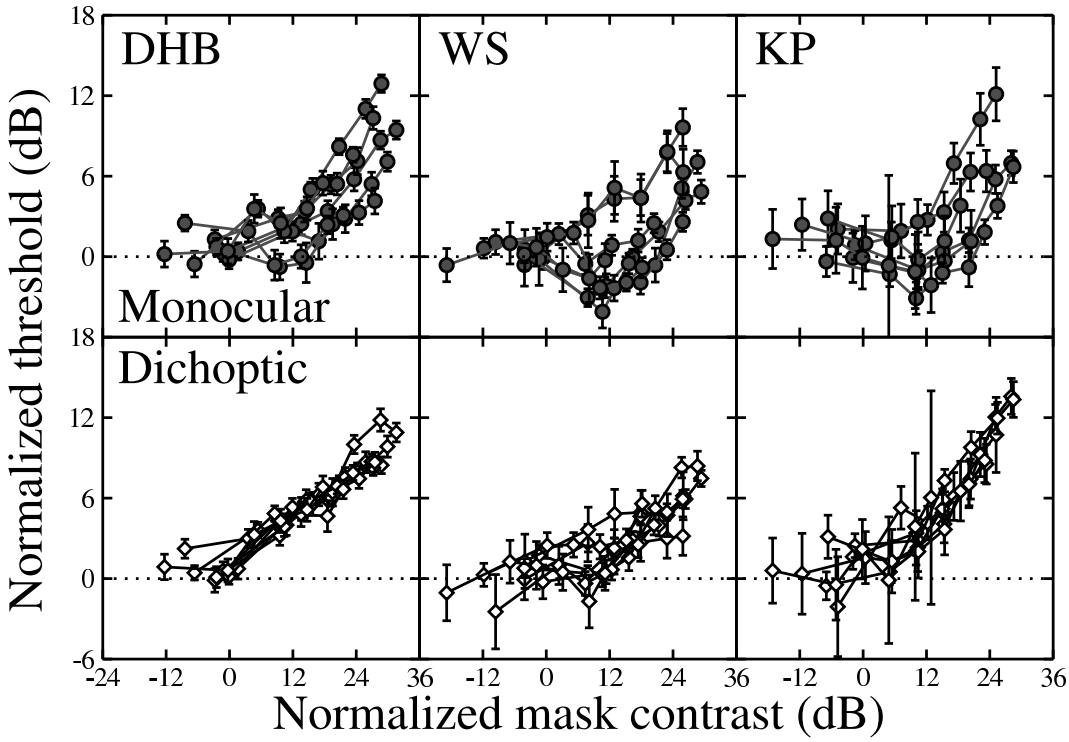

Figure 5.16: Threshold normalized masking functions, collapsed across spatiotemporal condition. Upper panels show monocular data, and lower panels show dichoptic data. For each function, the data have been normalized to detection threshold along both axes. Thus, thresholds below the horizontal zero line indicate facilitation, and those above it, masking. Mask contrasts <0dB are sub-threshold, and those >0dB are suprathreshold.

5.6.3 Modelling

Initially, the model used by Meese and Holmes (2007) was fit to the data. However, this provided a poor description of the dichoptic data, for reasons further discussed below. Instead, the preferred model from Experiment 7 was used to fit the data, using a procedure similar to that of Meese and Holmes (2007). In order to capture the monocular facilitation found for subjects WS and KP, a facilitation term was added to stage two, as described in Chapter 4. The model equations for the test channel are,

$$r_{iL} = \frac{C_{iL}}{1 + \omega_M C_{jL}}, \qquad r_{iR} = \frac{C_{iR}}{1 + \omega_M C_{jR}}, \tag{5.11}$$

$$r'_{iL} = \frac{{r_{iL}}^m}{S + r_{iL} + r_{iR}}, \qquad r'_{iR} = \frac{{r_{iR}}^m}{S + r_{iR} + r_{iL}}, \tag{5.12}$$

$$r''_{iL} = \frac{r'_{iL}}{1 + \omega_D r'_{jR}}, \qquad r''_{iR} = \frac{r'_{iR}}{1 + \omega_D r'_{jL}}, \tag{5.13}$$

$$resp = \frac{(r''_{iL} + r''_{iR})^p (1 + \alpha(r''_{jL} + r''_{jR}))}{Z + (r''_{iL} + r''_{iR})^q}, \tag{5.14}$$

with all parameters as described previously. The parameters from Meese *et al.* (2006) were again used to constrain the model. Since the model was fit to normalized data, the model responses were also normalized to the predicted detection threshold (as were the mask contrasts). Figure 5.17 shows the present form of the model, which includes monocular cross-channel suppression before stage one, dichoptic cross-channel suppression after stage one, and cross-channel facilitation at stage two.

Initially, the weight of facilitation (α) was fixed at zero. A sequential fitting procedure was then used to find the best estimates of the weight (ω) parameters. This was possible because varying a weight parameter for one condition affected only the data points in that condition. As there were eight spatiotemporal conditions, and two ocular conditions, there were 16 free model weights per subject (14 for KP), or one free parameter per masking function. Using a standard multidimensional simplex algorithm with this number of free parameters is cumbersome, and prone to errors from local minima. Sequential fitting, on the other hand, is much faster, and consistently produces good fits to each function.

In order to determine the optimal weight of facilitation, the fitting was repeated with a range of α values. Figure 5.18 shows how the RMS error across the entire data set varied with α. A clear nonzero minima is apparent for both KP and WS, the observers who display

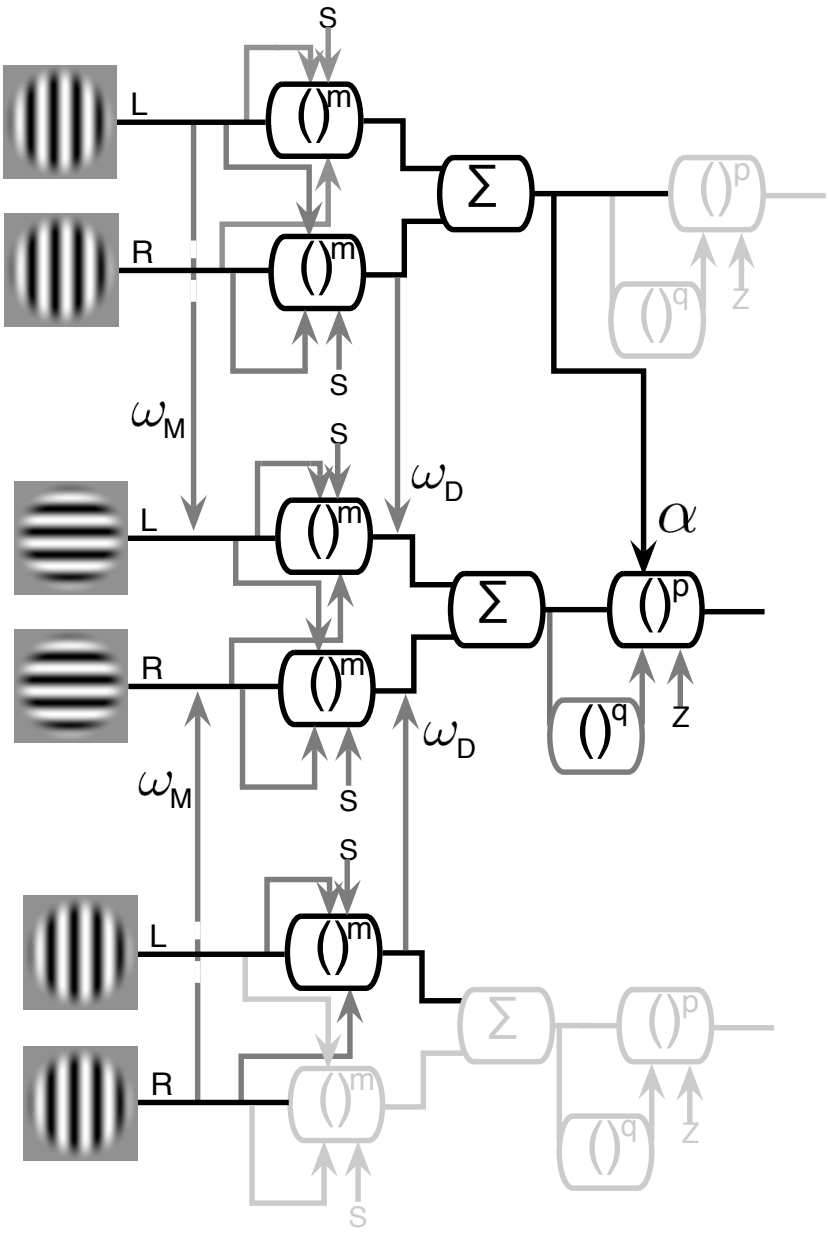

Figure 5.17: Diagram of cross-channel masking model. The format follows that of Figure 5.8, in that the central system is sensitive to the horizontal test stimulus, and the flanking systems are identical mechanisms sensitive to the mask. Grey arrows indicate divisive suppression, which can be weighted by model parameters (i.e. ω_M). Facilitation occurs at stage two, and is weighted by the parameter α. Note that arrows in which the test mechanism suppresses the mask mechanisms are omitted for clarity, but are still implemented computationally.

cross-channel facilitation. For subject DHB, who does not, the best fitting value by a small margin is α=0.1. The best fits to the normalized data are shown in Figure 5.19, using these optimal α values. Overall, the fits are very good, with RMS errors of <1.3dB for each subject (DHB=0.81dB, WS=1.26dB, KP=1.22dB).

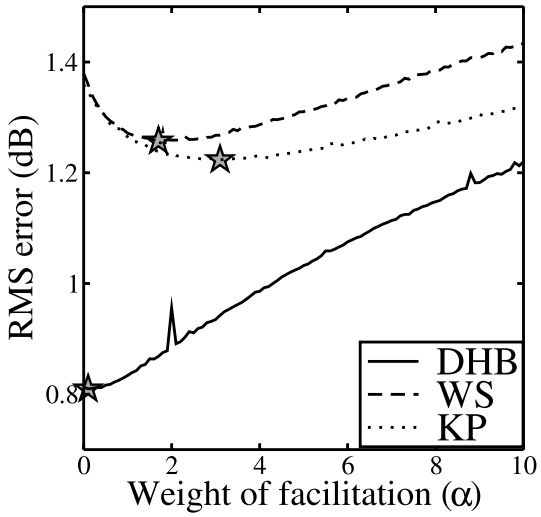

Figure 5.18: Change in RMS error with α. Functions were sampled with a resolution of 0.1. For each α value, all model weights were set by the sequential fitting procedure described in the text. The two subjects who show cross-channel facilitation have a nonzero minima in their functions, marked by a star (WS α=1.7, KP α=3.1). The function for subject DHB, who showed no facilitation, has a minima at α=0.1.

The weight parameters are displayed graphically in Figure 5.20, and show a consistent trend across subjects. On log axes, the monocular weights increase linearly as a function of speed. Best fitting regression lines have a mean slope (exponent) of 0.51, indicating that the weight is determined by $\sim\sqrt{TF/SF}$. This is the same result as described by Meese and Holmes (2007) for binocular presentation. The dichoptic weights, on the other hand, show almost no speed dependency. They are constant across all speed conditions, with best fitting regression slopes close to 0. This supports the conclusion that dichoptic cross-channel masking is spatiotemporally scale invariant.

5.6.4 Discussion

Cross-channel masking was measured over a wide range of spatiotemporal frequencies, for both monocular and dichoptic presentation. Monocular masking varied with the square root of stimulus speed, and some facilitation was observed. Dichoptic presentation produced no facilitation, and masking was equal in all spatiotemporal conditions (on normalized axes). Dichoptic cross-channel masking was thus found to be scale invariant.

The similarity between the monocular results here and the binocular results of Meese and Holmes (2007) is reassuring. As noted previously, monocular and binocular masking have

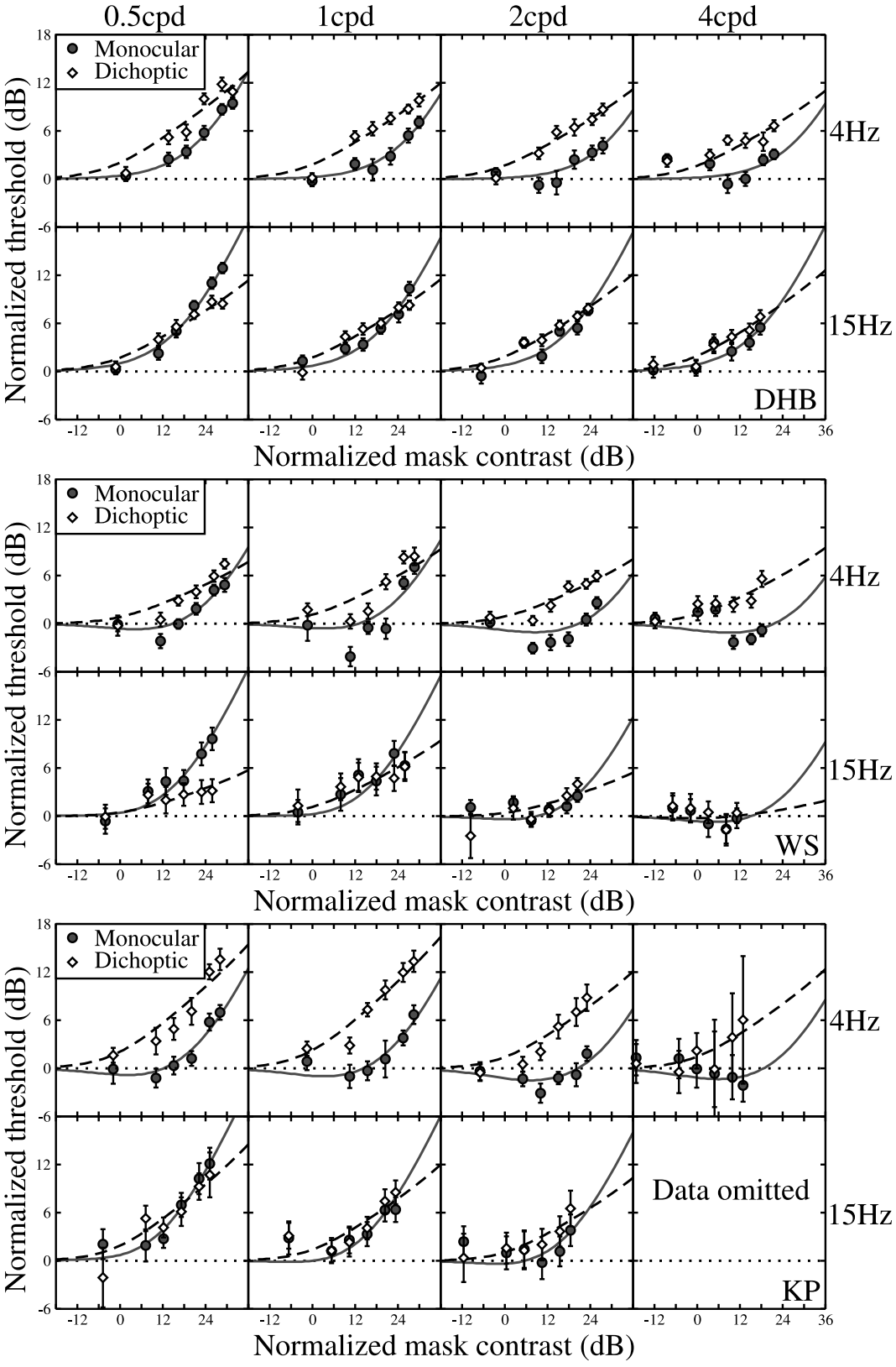

Figure 5.19: Model fits to normalized data. Panels have the same layout as for Figure 5.15. Curves are model fits to the monocular (grey continuous lines) and dichoptic (black dashed lines) data, normalized to detection threshold. The data for subject KP at 4cpd, 15Hz were omitted for reasons described in the text. RMS errors were below 1.3dB, using the parameters depicted graphically in Figures 5.18 and 5.20.

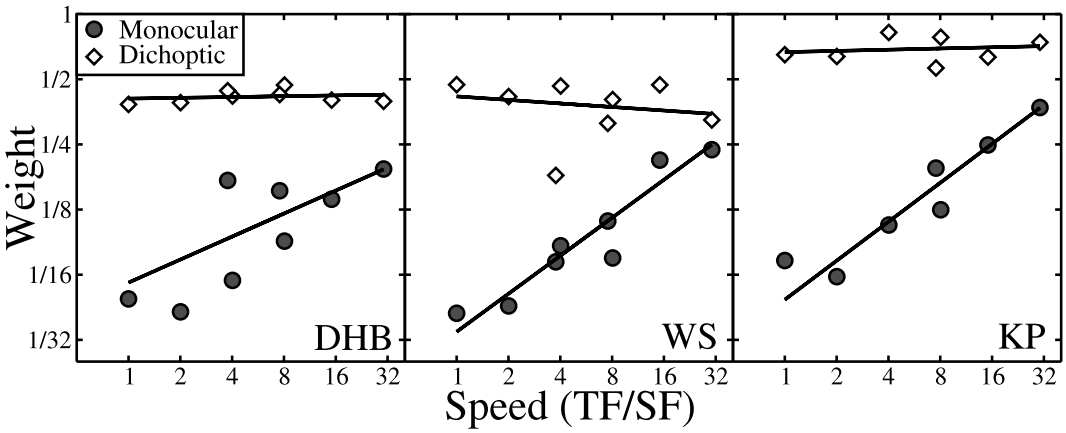

Figure 5.20: Best fitting model weights for monocular and dichoptic conditions, plotted against speed. Straight lines are best fitting log-regressions, calculated by finding optimal parameters for the function $y = mx^c$. The monocular weights increase in approximate proportion to $\sqrt{TF/SF}$ (exponents of best fitting regressions: DHB; $c = 0.35$, WS; $c = 0.58$, KP; $c = 0.60$). The dichoptic weights show no speed dependence, with regression exponents close to 0 (DHB; $c = 0.01$, WS; $c =$ -0.05, KP; $c = 0.02$). The fits themselves are shown in Figure 5.19, and are all good.

similar properties, as one would expect from the observation that the world looks the same when viewed with either one or two eyes. Meese and Holmes (2007) point out that since fast stimuli preferentially excite the magnocellular pathway, m-cells might be responsible for the greater masking effects observed at high speeds. This is consistent with the finding that m-cells are quite nonlinear, and saturate at high contrasts (i.e. Derrington and Lennie, 1984). The purpose of cross-channel suppression could be to provide normalization to m-cells and allow them to operate over a greater dynamic range. The parvocellular pathway is almost linear, so p-cells would not require such normalization.

The findings from single-cell physiology described in Section 5.1, suggest a similar explanation. Li *et al.* (2006) present convincing evidence that monocular cross-channel masking is *caused* by nonlinearities in LGN neurones (see also, Priebe and Ferster, 2006). These nonlinearities occur within-channel in isotropic LGN receptive fields, and include saturation (compression) and thresholding. Thus, m-cells would be expected to produce the greater masking, as they are more nonlinear than p-cells. These two accounts offer at least some insight into the speed relationship, although it is surprising that the behavioural manifestation (i.e. masking) of these physiological effects (i.e. saturation) should prove to be so lawful.

The finding that dichoptic masking is spatiotemporally scale invariant was unexpected. In one sense, it is not surprising that the dichoptic results are very different from the monocular data, given that the two types of masking are believed to have distinct neural loci. On the other hand, such consistency over a wide range of stimulus values (5 octaves of speed) suggests that scale invariance is a fundamental property of suppression in cortex. If so,

it could have relevance for other aspects of cortical suppression, such as binocular rivalry (i.e. Alais and Blake, 2005), surround suppression (Meese *et al.*, 2007b; Bonin *et al.*, 2005), and interocular suppression in amblyopia (Harrad and Hess, 1992; Levi *et al.*, 1980, see also Chapter 7).

Support for this hypothesis comes from a parallel study, which further extended the Meese and Holmes (2007) paradigm, and reported scale-invariant masking for binocularly presented chromatic (isoluminant) stimuli (Medina, Meese and Mullen, 2007). This study is discussed in greater detail in section 8.5.2. Research into motion perception has also revealed self-normalizing (scale invariant) processes, suggesting that these might be important in the cortical areas devoted to motion (Rainville, Makous and Scott-Samuel, 2005). Finally, there is also some evidence of individual differences in the level of dichoptic masking (for example, compare dichoptic data for WS and KP in Figure 5.16), which are discussed further in section 8.3.2. A direct, within-subjects comparison of the individual differences in dichoptic masking with those reported in other paradigms, such as area summation (Meese, Hess and Williams, 2005b) and binocular rivalry (Pettigrew and Miller, 1998), would provide a good test of the idea that there is a common factor underlying several forms of cortical suppression.

Since the dichoptic data collapse onto a single function (Figure 5.16), a single weight can be fit to the entire set of dichoptic data for each subject. This has two advantages. Firstly, it allows thresholds to be predicted for any relative mask contrast in any spatiotemporal condition, not just the ones explored in the above experiment. Secondly, goodness of fit for two possible dichoptic models can be compared.

In the model described by equations 5.11-5.14, the dichoptic mask signal has passed through stage one of the model before it impacts the test channel. This arrangement was initially introduced in Chapter 4 in order to encompass the half-binocular paradox (in which cross-channel dichoptic masking is reduced by the introduction of an additional monocular mask). A consequence of this is that the mask signal is normalized by stage one, producing a straight masking function on log axes. This is substantially different from the exponential-type function seen when i) the input contrast is used as the mask signal in the model, and ii) the mask is presented monocularly (experimentally). Up until this point, there has been no formal assessment of the validity of normalizing dichoptic signals in this way.

Two models were fit to the threshold-normalized dichoptic data from Experiment 10 and their performance compared. Model 1 used only the dichoptic mask contrast as a suppressive term impacting after stage one, such that equation 5.13 becomes,

$$r''_{iL} = \frac{r'_{iL}}{1 + \omega_D C_{jR}}, \qquad r''_{iR} = \frac{r'_{iR}}{1 + \omega_D C_{jL}}, \tag{5.15}$$

		Model 1		Model 2
Subject	ω_D	RMS Error (dB)	ω_D	RMS Error (dB)
DHB	0.056	2.24	0.40	0.86
WS	0.041	1.44	0.24	1.49
KP	0.085	1.47	0.45	1.75

Table 5.3: Parameters and RMS errors of two models for the dichoptic data of Experiment 10. Models are as described in the text. Errors are RMS errors in dB. The best fits are shown in Figure 5.21.

with all terms as previously. Model 2 used the normalized contrast signal (output of stage one), impacting at the same site, as described by equation 5.13. This is the model used throughout this chapter, constrained such that the dichoptic weight must be the same for all spatiotemporal conditions[6]. Each fit required only a single free parameter; the weight of dichoptic masking, ω_D. RMS errors and parameters are given in Table 5.3 for both models, with the fits shown in Figure 5.21.

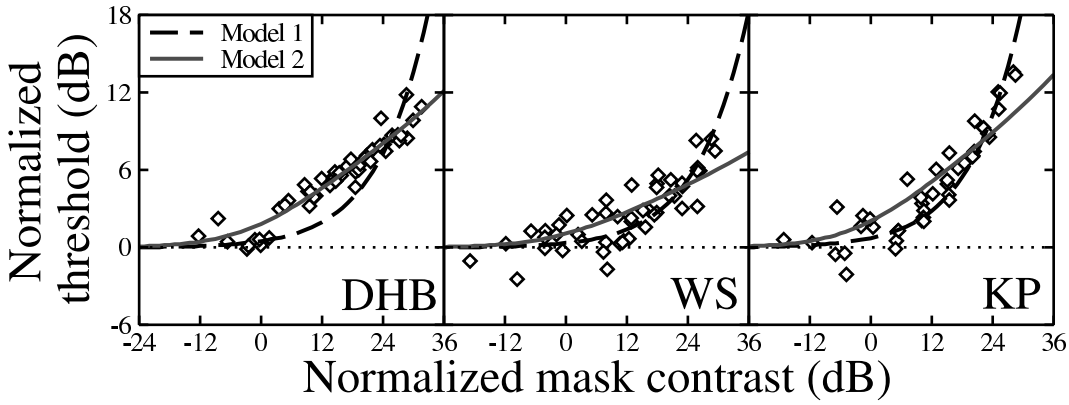

Figure 5.21: Best fits of two models to the threshold normalized dichoptic data of Experiment 10. Data points are replotted from Figure 5.16, with error bars omitted for clarity. Curves are the best fits of two models to the entire data set, as described in the text, with model parameters and RMS errors given in Table 5.3.

The two models make substantially different predictions. As expected, Model 1 produces an exponential type masking function, whereas Model 2 produces a straight masking function. For DHB, Model 1 fits very poorly indeed, with an RMS error 1.38dB greater than that of Model 2. It falls far from most of the data points by a wide margin. For the other two subjects, Model 1 is either equivalent (WS) or slightly better (KP) than either of the other two models when RMS errors are compared.

Thus it cannot be concluded that normalization of the dichoptic signal always produces a better fit, although some data do support this. It should be noted that the data for subject DHB have smaller errors and are generally less variable, due to this observer being much more psychophysically experienced than either of the other two. The poor fit of Model

[6]This model produced equivalent predictions with and without facilitation. The fits shown are for α=0.

1 to the data of DHB also explains why the model of Meese and Holmes (2007) was not appropriate for these data. In that model, masking is also a function of mask contrast, and produces exponential-like masking functions.

The scale invariance of dichoptic masking means that the only determinant of the strength of masking is the contrast of the mask, relative to threshold. This is consistent with the dependency on stimulus duration observed in the experiments of Chapter 4. There, the dichoptic data were modelled well by using the same function of duration as was used to describe detection thresholds (see also, Baker *et al.*, 2007c). However, when considered in terms of relative threshold elevation, all of the dichoptic data collapse approximately onto a single function, independent of duration, as shown in Figure 5.22 for subject DHB.

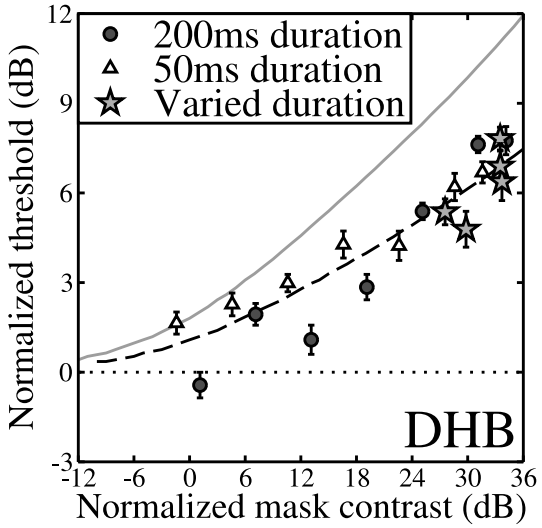

Figure 5.22: Dichoptic data from Chapter 4 (1cpd, orthogonal mask), normalized to detection threshold. The data fall approximately on a single function, independent of duration, particularly at the high mask contrasts. This function is shallower than that derived from the Experiment 10 data (grey curve). When the model is refit to the Chapter 4 data, ω_D changes from 0.4 to 0.24, and the fit improves dramatically (RMS errors: 3.26dB (grey curve), 0.87dB (black dashed curve)).

Interestingly, this dichoptic function is shallower than that derived above (compare data to grey curve). This is most likely to be due to differences in the stimuli between experiments: in Chapter 4, grating stimuli with raised cosine envelopes were used, whereas Experiment 10 here used gabor functions. The temporal waveforms also differed between experiments, as Chapter 4 used static stimuli, whereas those in Experiment 10 were temporally modulated (recall that modulated stimuli also produced stronger masking in Experiment 9). The model can provide a good fit to these data also by re-estimating the dichoptic weight (ω_D; black dashed curve). This indicates that although the scale invariance finding holds within a set of stimuli, there might be differences as stimulus size and temporal profile are varied.

5.7 Chapter Summary

Inspired largely by recent neurophysiological findings, four experiments were devised to explore cross-channel masking for monocular and dichoptic presentation. Monocular masking was found to be i) undiminished after adapting to the mask, ii) broadly tuned, and iii) dependent on stimulus speed. Dichoptic masking was found to be i) reduced after adapting to the mask, ii) less broadly tuned than monocular masking, and iii) spatiotemporally scale invariant.

All of these findings strongly support the conclusion from Chapter 4 that monocular and dichoptic masking are caused by different suppressive processes. The neurophysiology places monocular suppression in the LGN, and dichoptic suppression in visual cortex. This is consistent with the architecture of the preferred psychophysical model, based on the results of Experiments 7, 8 and 10, and the conclusions of Chapter 4. The model successfully accounts for a wide range of experimental data, and is reasonably biologically plausible, despite some generalisations to retain simplicity.

Chapter 6

Experiments using stimuli in dichoptic antiphase

6.1 Experiment 11: dichoptic masking with antiphase stimuli

6.1.1 Introduction

In the Two Stage model, the strong masking caused by a dichoptic pedestal (Weber's law; see Chapter 3) occurs because of two processes. The first, termed the *direct effect*, is due to divisive interocular suppression of the test by the pedestal, and occurs at stage one. The second process, the *indirect effect*, is due to suppression of the pedestal by the test, and their subsequent combination at the binocular summation stage. When the dichoptic pedestal contrast is high, introducing a test signal can reduce the overall model output by an amount greater than it contributes. This is the effect which produces the paradoxical psychometric functions for high contrast dichoptic masks explored in section 3.4. It also raises thresholds, as the test contrast must be further increased in order to overcome the indirect effect.

The presence of two distinct processes in the model is a novel and interesting method for producing Weber's law behaviour. However, there is no strong evidence to favour this arrangement. What is required is some method of isolating the two effects, for example by circumventing binocular summation. One possibility is to present stimuli 180° out of phase (in antiphase) across the eyes. As will be discussed below, stimuli in this configuration apparently do not undergo binocular summation.

At detection threshold, stimuli in dichoptic antiphase can be detected (Legge, 1984a; Simmons, 2005). This suggests that the brain does not combine the signals algebraically, as

has been suggested (Ding and Sperling, 2006b), since this would result in cancellation. The summation ratio for antiphase stimuli (when compared to monocular detection of only one grating) is typically between 1 and 1.2 (Bacon, 1976; Simmons, 2005). This is weaker than the summation observed between stimuli which are in phase in the two eyes (see section 1.4.1) and is of the magnitude expected from probability summation of two independent noisy signals (Pirenne, 1943; Eriksen, 1966; Tyler and Chen, 2000). Similar findings have been reported by Green and Blake (1981), using sequential dichoptic presentation of pairs of gratings, either in-phase, or in antiphase, and by Westendorf and Fox (1974) using flashes of light.

These findings indicate that antiphase stimuli are not subject to the neural binocular summation which occurs for in-phase gratings. This suggests that they may offer a useful tool for studying the mechanisms of interocular suppression in isolation, without contamination from the indirect effect (which relies on binocular summation). Antiphase stimuli have not previously been used in a dichoptic masking paradigm, so nothing is known about their suppressive effects. However, given the substantial dichoptic masking found using stimuli of different orientations and spatial frequencies (see Chapters 4 and 5), it seems likely that interocular suppression will not be strongly phase dependent.

That dichoptic masking may be phase dependent is also consistent with single cell data from cat. Ohzawa and Freeman (1994) and Truchard, Ohzawa and Freeman (2000) measured the responses of binocular V1 cells to interocular pairs of gratings of different phases and contrasts. They found that firing was reduced at mismatched phases, and minimal for antiphase stimuli. However, this result was attributed to linear filtering (i.e. cancellation) of signals, which is not consistent with the psychophysical results discussed above (Legge, 1984a; Simmons, 2005). It is not clear how these conflicting findings can be resolved.

Experiment 11 repeats the dichoptic masking condition of Experiment 1 (Chapter 3), using gratings which are either in phase, or out of phase by 180°.

6.1.2 Methods

Apparatus and Stimuli

Horizontal gratings at 1cpd were presented using a mirror stereoscope. A ViSaGe was used to display the stimuli on a Clinton Monoray monitor. All other methodological details were as described in Chapter 2.

Procedure

Dichoptic masking using in-phase stimuli is a contrast discrimination task, in which the observer indicates which interval contained the highest contrast grating. However, stimuli in antiphase are likely to require a different strategy, more akin to contrast detection, since the mask and test will be detected by different phase channels. For this reason, the two conditions were blocked separately. Two interleaved '3-down, 1-up' staircases recorded performance in the left and right eyes. Conditions were blocked by mask contrast, which ranged from 0-30dB in steps of 5dB, and also included a 33dB mask, and a 0% mask (detection threshold). Each subject repeated the experiment 4 times, and the data were pooled across eye of presentation and session, before using probit analysis to estimate a single threshold.

As described for previous experiments, the absolute phase of the test was randomly selected on each trial from four possible phases. Thus, the important factor was the *relative* phase between the test and the mask, which was either 0° (in-phase) or 180° (antiphase).

Observers

LP was a 24 year old female undergraduate student, who was psychophysically naïve, and was not aware of the aims of the experiment. Both subjects (LP and DHB) were emmetropic, and had no abnormalities of binocular vision.

6.1.3 Results

Figure 6.1 shows the results for both subjects. The in-phase data are in good agreement with the findings reported in Experiment 1 (Chapter 3). A shallow region of facilitation is apparent at mask low contrasts, followed by strong masking at high contrasts, which approximates Weber's law (slope of ∼1). The antiphase results, however, are clearly different. There is no evidence of facilitation at low contrasts[1], and the masking function is shallower (slope of ∼0.6). Weaker masking is therefore observed at the higher contrasts, by as much as 12dB (a factor of 4).

6.1.4 Modelling - the nature of dichoptic masking

When dichoptic masks are presented in antiphase to the test, they produce weaker masking than equivalent masks which are in-phase. It can be determined computationally whether this is by an amount consistent with the removal of binocular summation from the model.

[1]Curiously, Blake and Levinson (1977) do report dichoptic facilitation using (vertical) antiphase stimuli. However, as Green and Blake (1981) later point out, this is most likely attributable to vertical misalignment of the stimuli, caused by ocular vergence movements.

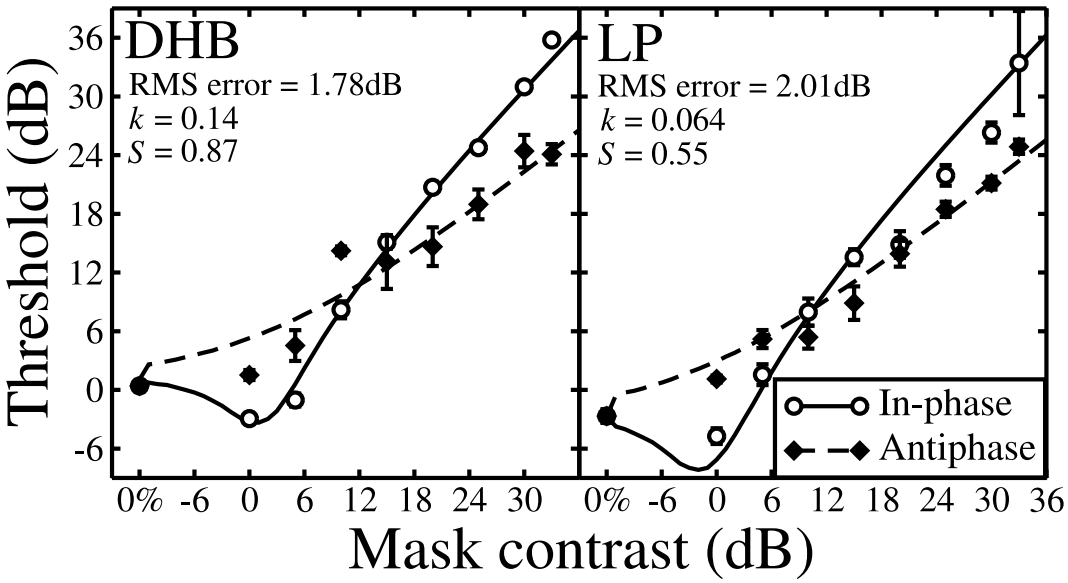

Figure 6.1: Dichoptic masking functions with mask and test either in-phase (open symbols), or out of phase by 180° (filled symbols). In-phase data replicate those reported in similar experiments (i.e. Experiment 1 (Chapter 3); Meese *et al.*, 2006), whereas antiphase masks produce weaker masking and no facilitation. Curves are predictions of the Two Stage model, as described in the text. Error bars give the standard error of the probit fit.

Removing binocular summation will destroy dichoptic facilitation, as this is dependent on combination of mask and test prior to the second gain control stage of the model. It will also reduce the level of masking, as the contribution from the indirect effect (see section 6.1.1 above) will not occur. Thus, the model without binocular summation will produce a monotonic masking function with a slope shallower than that of the in-phase function, as found empirically.

The model was extended to include four identical mechanisms, tuned to the four cardinal phases used in the experiment. Each mechanism inflicts interocular suppression on the others, so suppression is phase insensitive. However, binocular summation occurs only within a mechanism. Stage one of each phase mechanism is expressed as,

$$LE_{\theta_0} = \frac{C^m_{L\theta_0}}{S + C_{L\theta_0} + [C_{R\theta_0} + C_{R\theta_{90}} + C_{R\theta_{180}} + C_{R\theta_{270}}]}, \tag{6.1}$$

where θ indicates preferred phase by its numeric subscript, and all other terms are as used previously. The square brackets enclose the interocular suppressive terms to aid the reader's segmentation of the equation, and correspond to a dichoptic 'gain pool' (Heeger, 1992). An equivalent expression exists for the right eye, and for the other three phase mechanisms. Each phase mechanism then undergoes binocular summation and stage two of the model, determined by,

$$resp_{\theta_0} = \frac{(LE_{\theta_0} + RE_{\theta_0})^p}{Z + (LE_{\theta_0} + RE_{\theta_0})^q}, \tag{6.2}$$

with parameters retaining their meanings from previous chapters. The output stage of the model is a MAX operator which selects the phase mechanism with the largest increase in activation between the null and test intervals. This operation reflects an observer's ability to select the channel containing the additional (test) stimulus, and is distinct from the traditional implementation of a MAX operator, which selects the channel with the largest output in each interval. This traditional MAX operator predicts that test contrast must exceed mask contrast to be detected, which is clearly not the case here as it predicts a masking function with a slope of unity. Thus, the MAX operator which selects the greatest increase between intervals is a mathematical convenience, approximating a more complex aspect of observer behaviour. Figure 6.2 shows the model architecture diagrammatically for a single phase mechanism.

Assessed in isolation, each phase mechanism is equivalent to the standard Two Stage model, with no additional parameters. This means that many of the parameters can be fixed at values derived previously by Meese *et al.* (2006). The four fixed parameters were: $p = 7.99$, $q = 6.59$, $m = 1.28$, $Z = 0.076$. The two remaining parameters, k and S, were adjusted using a simplex algorithm to provide the best possible fit to the in-phase data for each subject (Figure 6.1, solid curves, parameter values given in figure). The antiphase results were then predicted with all parameters fixed (dashed curves). RMS errors given in the panels are calculated across both conditions.

The model predicts the antiphase data well. As it was not necessary to change the weight of interocular suppression across phase conditions, the implication is that the dichoptic suppressive mechanisms of stage one are phase insensitive. Since the antiphase data are consistent with a version of the model in which binocular summation does not occur between phase mechanisms, they indicate that binocular summation is phase dependent. This provides strong evidence that there are indeed two distinct processes responsible for dichoptic pedestal masking, and that stimuli presented in dichoptic antiphase provide a tool for studying their properties.

6.1.5 Phase tuning of dichoptic masking

The results of a supplementary experiment, carried out at McGill vision research unit, using shutter goggles and a VSG 2/5, are shown in Figure 6.3. Here, the phase relationship between mask and test was varied in steps of 15°. Data were gathered at two mask contrasts, 30dB and 5dB, to illustrate the transition between the two magnitudes of masking at high mask

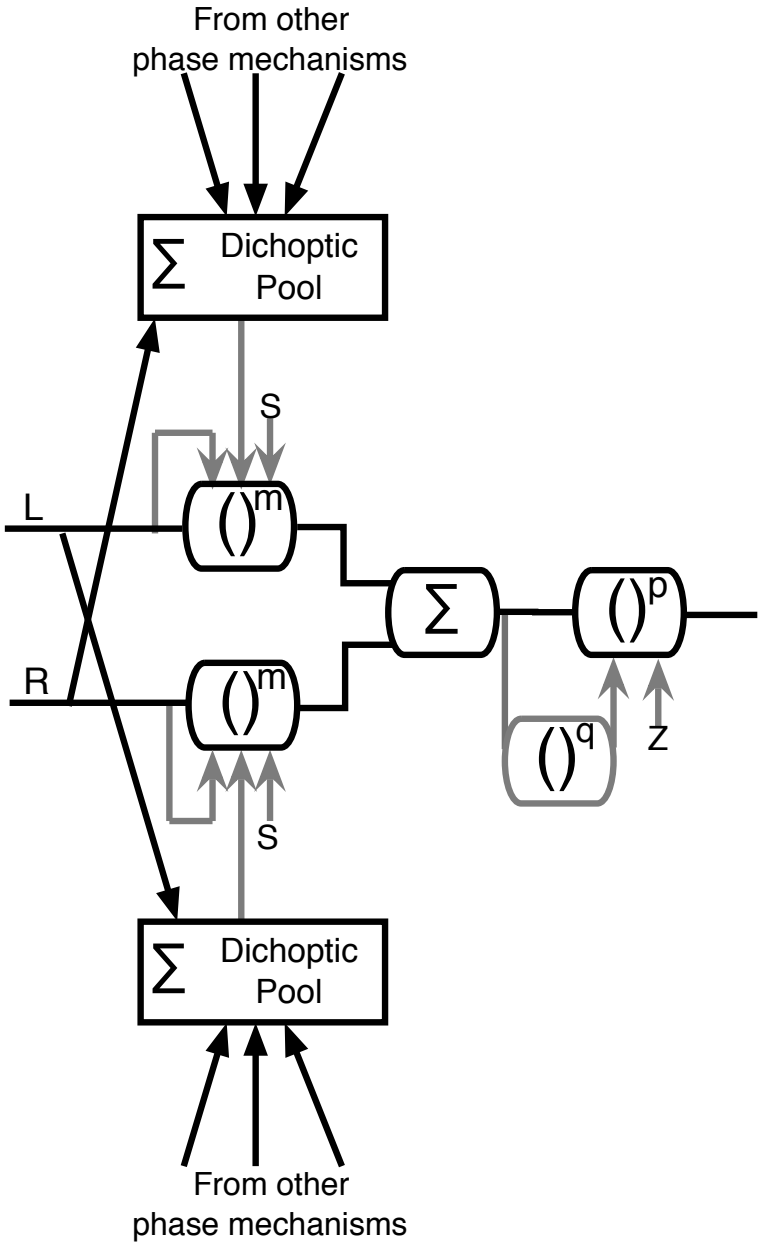

Figure 6.2: Diagram of dichoptic phase model. Depicted is a single mechanism, optimally sensitive to a specific phase (i.e. 90°). Dichoptic suppression occurs across mechanisms sensitive to all phases, governed by the dichoptic pool, which corresponds to the terms in square brackets in equation 6.1. As for previous model diagrams, black arrows indicate excitation, grey arrows indicate divisive inhibition, Σ indicates summation, and brackets raised to a power denote exponentiation.

contrasts, and between masking and facilitation at low mask contrasts. The experiment was repeated six times, with each data point representing around 550 trials.

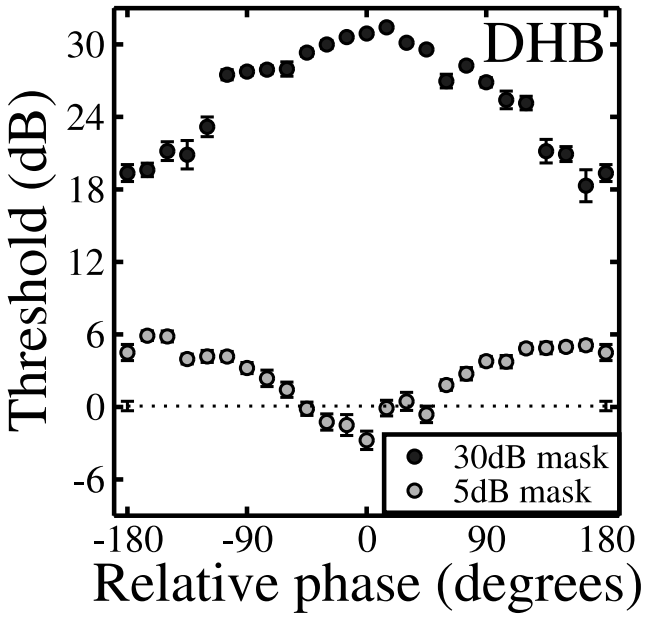

Figure 6.3: Phase tuning of dichoptic masking. Each data point is based on 6 repetitions (550 trials on average). The dotted line gives the baseline detection threshold (0% mask), and symbols show thresholds for a 30dB (dark fill) or a 5dB (light fill) dichoptic mask. Phase is reported relative to that of the test grating. The points at 0 and $\pm 180^\circ$ correspond to in-phase and antiphase conditions, respectively, in Figure 6.1.

The data show a smooth transition between the limiting conditions reported in the main experiment. This suggests that at intermediate phase differences (between 0 and 180°), partial binocular summation may occur. For the 5dB mask (light filled symbols), changes in threshold are only apparent for relative phases within $\pm 90^\circ$, whereas the 30dB mask (dark filled symbols) shows a continuous change in masking across all phases. Thus, it is possible that the phase tuning of binocular combination is contrast dependent.

6.2 Experiment 12: dichoptic spatial frequency tuning

6.2.1 Introduction - Legge's unusual masking functions

In his seminal study of dichoptic masking, Legge (1979) measured the spatial frequency tuning of dichoptic masking by varying mask spatial frequency over a wide range. The results at two test spatial frequencies (1 and 4cpd) are replotted in Figure 6.4, and reveal functions with an unusual shape. They peak sharply around the test frequency, but appear to be quite broadly tuned at more distant frequencies. It is as though the functions are a product of two different masking processes, one more narrowly tuned than the other.

However, Legge offered little explanation for this phenomenon in his paper, and it has not been explored in detail by subsequent studies, although it has been replicated (Levi *et al.*, 1979, 1980; Harrad and Hess, 1992).

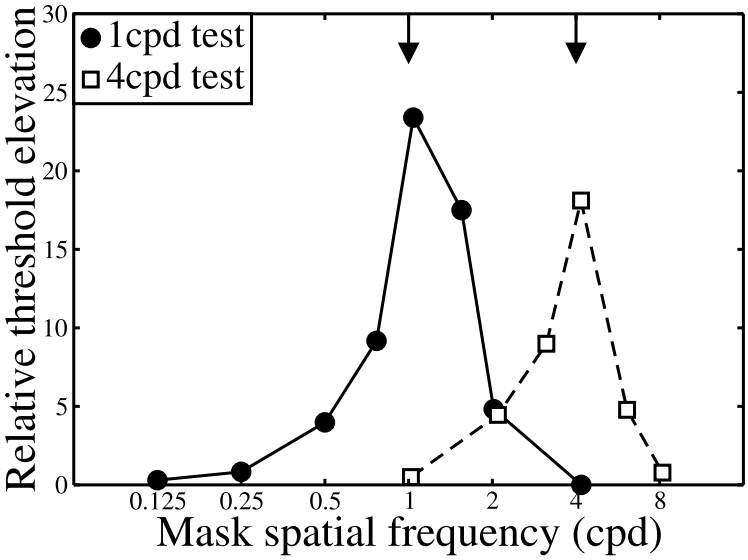

Figure 6.4: Replotted data from Figure 6 of Legge (1979). Test grating spatial frequencies are given by the arrows, and mask frequencies by the abscissa. The test was always presented to the right eye, and the mask to the left, simultaneously for a duration of 200ms. Mask contrast was 19% (26dB), and the test and mask were always in cosine phase with a central fixation point.

The findings of Experiment 11 suggest a possible explanation for Legge's data. The functions could be the envelope of broadly tuned interocular suppression (at stage one in the model) and narrowly tuned binocular summation. By repeating Legge's experiment using antiphase stimuli, it should be possible to map the tuning of interocular suppression, in the absence of binocular summation. This can then be compared with data for in-phase stimuli, to reveal the spatial frequency tuning of binocular summation (for these stimuli) as well.

6.2.2 Methods

Apparatus and Stimuli

Stimuli were generated at a range of spatial frequencies. It was necessary for all stimuli to be the same size, regardless of spatial frequency, because if the test were ever physically larger, it would be visible at the edges of the mask. Therefore, the same spatial window (a raised cosine envelope) was used for all stimuli.

Phase difference is always reported relative to the central fixation point, as for stimuli of different spatial frequencies, the phase difference is not necessarily constant across the

whole image. All apparatus was the same as for Experiment 11, except that subject LP used shutter goggles (CRS FE-1) instead of the stereoscope.

Procedure

Initially, masking was measured at nine mask spatial frequencies, ranging from 0.25 to 4cpd in logarithmic steps. The test spatial frequency was always 1cpd, and the mask contrast was always 30dB. All stimuli were horizontal, and all mask frequencies were interleaved within a block. In order to keep session duration reasonable, there were two blocks, each of which included only one staircase for each frequency, testing either the left or right eye. Using this design, blocks took around 15 minutes to complete. There were two blocks for the in-phase, and two for the antiphase condition, which were run in random order. Baseline detection thresholds (0% mask contrast) were also recorded. The experiment was repeated four times by each subject.

Observers

Subjects were the same as for Experiment 11. In the intervening months between experiments, subject LP had performed additional psychophysical experiments close to detection threshold. This practise appears to have resulted in an improvement in her baseline detection threshold, although it is possible that this is a result of using the goggles rather than the stereoscope. However, other subjects show negligible threshold differences between goggles and stereoscope (i.e. compare detection thresholds for subject DHB across Chapters 3 and 4), so this seems unlikely.

6.2.3 Results

As anticipated, the results for the antiphase condition, shown in Figure 6.5, are clearly different from those of the in-phase condition. Consistent with Legge's findings, the in-phase data have a sharp peak at the test spatial frequency. The antiphase condition does not peak as strongly, and appears to be maximal at a frequency just higher than the test frequency. It resembles a linear gaussian plotted on logarithmic axes (see modelling in section 6.4 below).

The results are strikingly similar across subjects. It is particularly interesting that the substantial difference in baseline detection thresholds (>6dB) appears to have little or no effect on the absolute vertical placement of the masking functions. It is possible that this represents some normalization process, much like the contrast constancy effect reported over different spatial frequencies and eccentricities (Cannon, 1985; Georgeson and Sullivan, 1975).

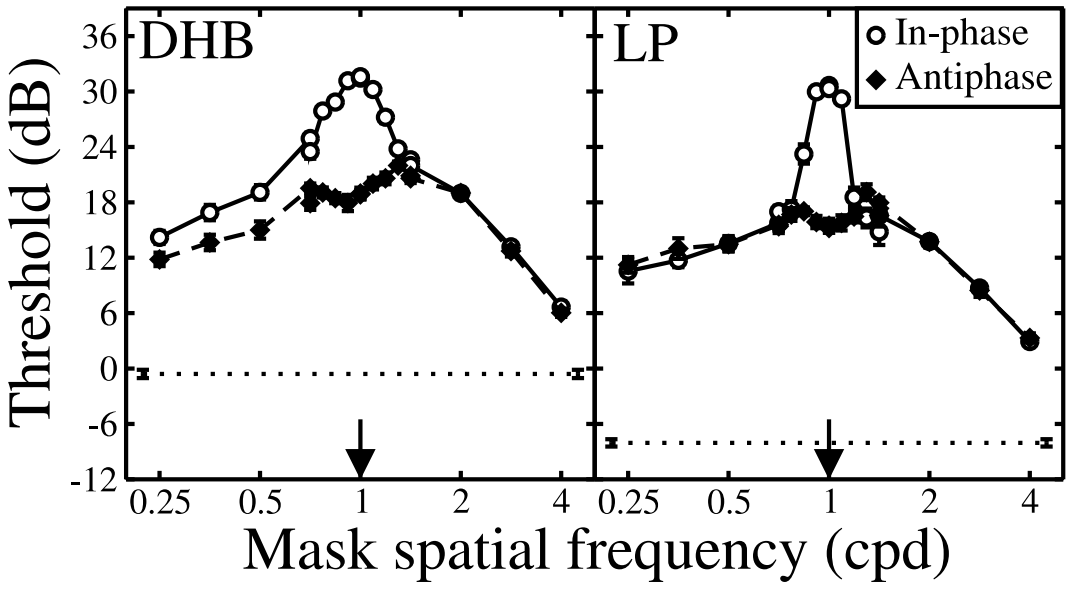

Figure 6.5: Dichoptic spatial frequency tuning functions for in-phase and antiphase gratings. All stimuli were horizontal, with test frequency indicated by the arrow. Open symbols show thresholds when mask and test had the same phase relative to a central fixation point. Filled symbols show results when mask and test were 180° out of phase. The horizontal dotted line is the baseline detection threshold (0% mask), and error bars give the standard error of the probit fit.

This is discussed further in Section 6.3.4.

The initial sampling regime ($\frac{1}{2}$ octave steps) lacked the resolution necessary to make strong conclusions about the spatial frequency tuning of binocular combination for these stimuli. The experiment was repeated, under the same conditions, but with much smaller increments in mask spatial frequency. The results are also shown in Figure 6.5, and reveal a gradual transition towards the peak for the in-phase condition for subject DHB. Subject LP shows a more abrupt transition, and slightly tighter tuning.

One clear difference between subjects is that the two functions do not superimpose at low mask frequencies for subject DHB, whereas they do for LP. Whether this represents a difference in attentional strategy, a difference between stereoscope and goggles, or a legitimate difference in the tuning characteristics of the subjects' visual systems is not clear. However, the finding was robust when the experiment was repeated with the finer sampling.

6.2.4 Discussion

Spatial frequency tuning functions for dichoptic masking were measured, where the mask and test were either in-phase or in antiphase. The in-phase data peaked strongly at the test frequency, showing around 12dB greater threshold elevation than the antiphase data. In combination with the approach taken for Experiment 11, these results suggest that the spatial frequency tuning of interocular suppression is broad, with a full width at half height

greater than two octaves, and a peak slightly above the test frequency. Binocular summation for these stimuli is much narrower ($< \frac{1}{2}$ octave), at least for frequencies higher than that of the test. More quantitative estimates of the tuning of these effects are derived below in section 6.4.2.

These findings provide further evidence that dichoptic masking consists of two separate processes, as suggested by the Two Stage model. Furthermore, this novel approach offers new insights into the unusual tuning functions reported by Legge (1979).

Binocular presentation methods

For these experiments, one observer (DHB) used the stereoscope, whilst the other (LP) used shutter goggles. These two presentation methods have their own relative merits and associated problems. With the stereoscope, there is absolutely no possibility of crosstalk (where an image intended for one eye is faintly visible to the other eye). This means that any interaction between the stimuli must come from the observer, and not from any physical interaction on the display. However, a small vertical misalignment of the mirrors could result in phase misalignments or vergence errors. This is not a problem with stereo goggles. It is reassuring that results are highly consistent using both presentation methods, indicating that none of these concerns impact on the experiment.

6.3 Experiment 13: dichoptic orientation tuning

6.3.1 Introduction

In principle, a similar approach to Experiment 12 can also be taken to orientation tuning. Orientation tuning functions for dichoptic masking have been reported before by Levi *et al.* (1979, 1980) and Harrad and Hess (1992), however these studies have several shortcomings. Both studies were primarily concerned with amblyopia, and so orientation tuning functions were gathered for only one normal observer in each case. Functions tended to be sparsely sampled, gathered at a higher spatial frequency than that used here (2 or 5cpd) with vertical stimuli. For the Levi *et al* experiment, mask contrast was only 10dB above threshold.

There is a need for orientation tuning data gathered at a high mask contrast, using low spatial frequency horizontal gratings (to avoid misregistration). This should enable characterisation of the orientation tuning of dichoptic suppression, particularly if the antiphase condition is also used. Experiment 12 was repeated using masks of the same spatial frequency as the test (1cpd), but different orientations.

6.3.2 Methods

All methods were as reported above, except that spatial frequency was kept constant (1cpd), and the orientation of the mask stimuli was varied from 0 to 165° (the test orientation was always 90°). Again, two sampling regimes were used: 3° increments close to the test orientation, and 15° increments for more extreme orientations.

6.3.3 Results

Figure 6.6 shows the orientation tuning functions for both subjects. As for Experiment 12, the in-phase condition has a sharp peak at the test orientation. The antiphase condition also peaks at the test orientation, but much less sharply. Indeed, the antiphase function is symmetrical about the test orientation, and resembles an almost perfect gaussian (see section 6.4). It appears that the orientation tuning function for binocular combination is also roughly gaussian, but has a much narrower bandwidth. As for spatial frequency tuning, subject DHB shows a smoother, broader function than LP.

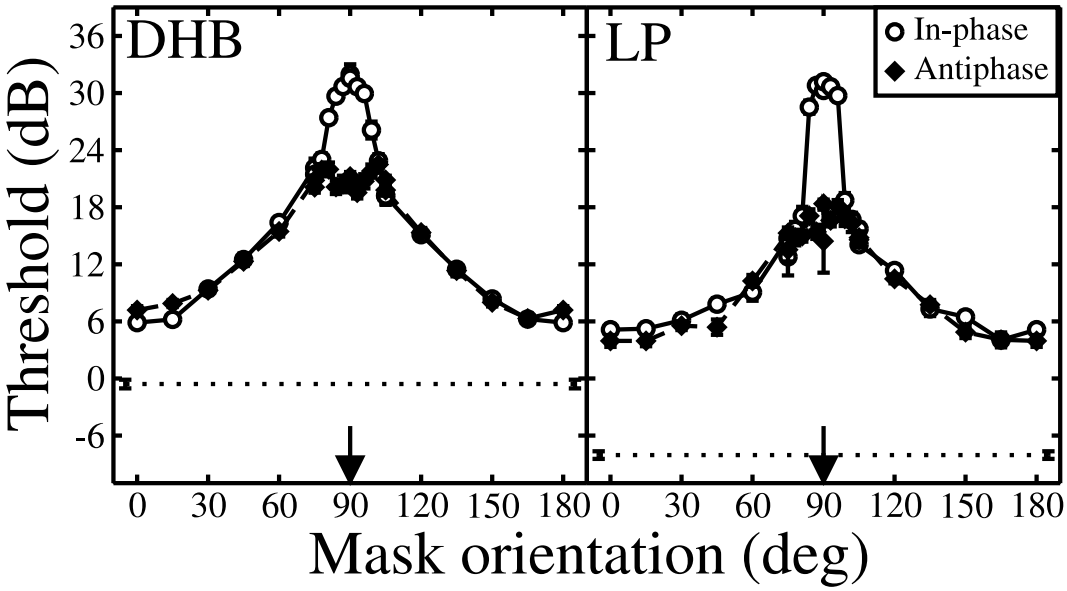

Figure 6.6: Dichoptic orientation tuning functions for in-phase and antiphase gratings. All stimuli were at 1cpd, with test orientation indicated by the arrow. Open symbols show thresholds when mask and test had the same phase relative to a central fixation point. Filled symbols show results when mask and test were 180° out of phase. The horizontal dotted line shows the baseline detection threshold (0% mask), and error bars give the standard error of the probit fit. The 0° and 180° conditions are equivalent, so these data were gathered only once, but plotted on both sides for symmetry.

Of particular interest are the subject differences when the mask is orthogonal to the test (orientation 0°). For these data, although the masked threshold is roughly the same for both subjects (∼6dB), the threshold elevation over the baseline is much greater for LP than

for DHB (12dB vs 6dB). This is discussed further below.

6.3.4 Discussion

Orientation tuning functions for dichoptic masking were measured, with stimuli either in-phase or in antiphase. Masking functions were symmetric about the test orientation, and resembled a gaussian, peaking most sharply when the stimuli were in-phase. The bandwidth of interocular suppression was broad, with masking evident for orthogonal masks, whereas the bandwidth of binocular combination was much narrower. Again, a more quantitative assessment of these bandwidths is given in section 6.4.

One key difference between subjects in Figure 6.6 is in the level of threshold elevation at 0° mask orientation. This condition is, of course, equivalent to the orthogonal masking conditions explored extensively in Chapters 4 and 5. In those experiments, substantial individual differences in the levels of dichoptic masking were found (see Chapter 8 for a direct comparison). Here also, the baseline detection threshold differed between DHB and LP by around 6dB, whereas the absolute vertical placement of the masking functions was the same for both subjects (with thresholds at 0° sitting at around 6dB). Thus, the contrast required to detect the test in the presence of an orthogonal mask appears to be constant, despite substantial differences in the baseline threshold.

With only two subjects in the orientation tuning experiment, it is not clear whether this is a robust finding, or merely coincidental. To test this, the experiment was repeated on two further subjects. Their results, shown in Figure 6.7, indicate that the finding is coincidental; thresholds for the orthogonal mask are around 15-18dB for both subjects. The baseline detection threshold again varies greatly with observer. These substantial subject difference in the level of dichoptic orthogonal masking are presumably due to individual differences in the suppressive weights of the underlying gain pool, and are discussed further in Chapter 8.

6.4 Modelling dichoptic masking functions

In this section the Two Stage model is extended in order to characterise the orientation and spatial frequency masking data. This will provide a quantitative summary of the results, and also allow the model to predict data from other conditions. Of particular interest for this purpose are estimates of the bandwidth of interocular suppression, which have not been reported previously. The bandwidth of binocular combination is also reported, although as discussed below this is likely to be stimulus specific due to the large stimulus size and phase

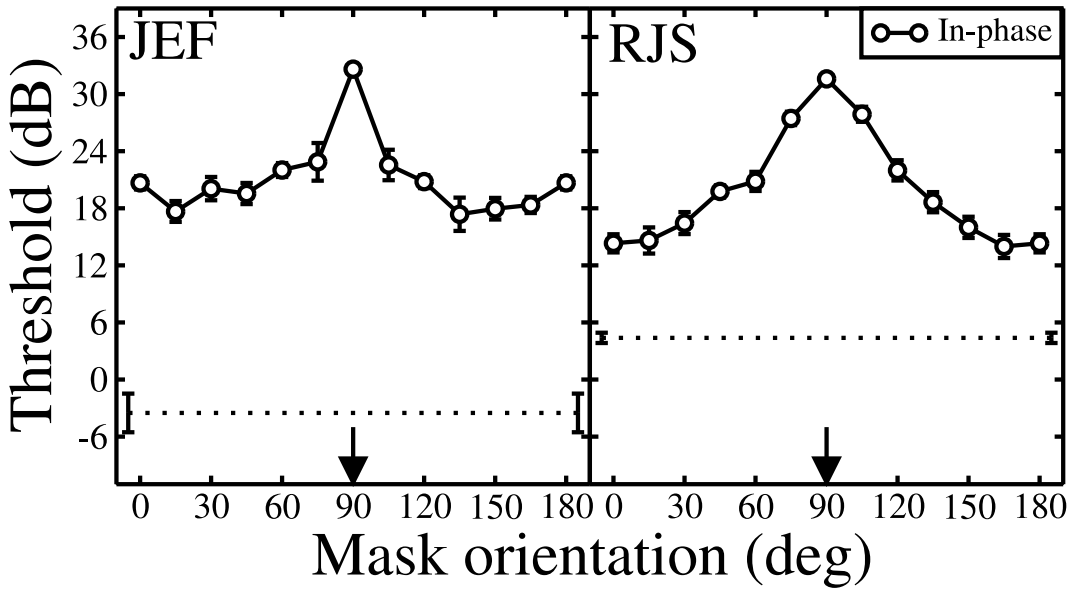

Figure 6.7: Further in-phase dichoptic orientation tuning functions. Details are as for Figure 6.6, with two different subjects, JEF (M/21) and RJS (M/30), who both show substantial orthogonal dichoptic masking. All stimuli were presented dichoptically using the mirror stereoscope.

specificity of the effect (see Bergen *et al.*, 1979).

Initially, the orientation tuning functions are modelled. This is because they are more straightforward than the spatial frequency functions, being both symmetrical about their midpoints, and also plotted on a linear abscissa. The model, once developed, is then easily applied to the spatial frequency data, with a few modifications.

6.4.1 Orientation model

For a given mechanism, interocular suppression can be well approximated by a gaussian function, plus an additive constant. The constant serves to shift the function vertically, producing threshold elevation even at the extremes of the function, as observed for the antiphase masking condition. This tuning function is expressed as,

$$\omega_{D(\phi)} = \frac{w + e^{\frac{-\phi^2}{2\sigma^2}}}{w+1}, \tag{6.3}$$

where w is a model weight, σ is the standard deviation of the gaussian, e is a constant, and ϕ is the difference (in degrees) between the mechanism's optimal orientation, and that of the stimulus in the other eye. This results in an elevated gaussian function, with a maximum value of 1, which occurs when the orientation of the mask equals the mechanism's optimal orientation. In other words, when mask and test have the same orientation, the model is equivalent to the standard Two Stage model, discussed in Chapter 3.

Subject	Figure	σ	ψ	w	S	k	RMS error (dB)
DHB	6.8A	22.9	3.37	0.05	1.51	0.0562	1.62
LP	6.8B	19.3	2.30	0.15	1.02	0.0006	1.44

Table 6.1: Parameters and RMS errors of the orientation model. Meanings of the parameter terms are as discussed in the text.

The model so far has considered mask stimuli which either excite the same detecting mechanism (i.e. pedestals), or stimuli which are sufficiently different that they do not excite the same mechanism (i.e. orthogonal masks) as the test. Here, the stimuli may partially excite a mechanism, contributing to its output by an amount proportional to that mechanism's sensitivity to such a non-optimal stimulus. This sensitivity function is described by a standard gaussian, without the need for an additive constant,

$$b(\phi) = e^{\frac{-\phi^2}{2\psi^2}}, \tag{6.4}$$

where terms are as for equation 6.3, except that the standard deviation of the gaussian here is termed ψ. Stage one of the model, including binocular summation, is therefore,

$$binsum = \frac{b(\phi_L){C_L}^m}{S + b(\phi_L)C_L + \omega_{D(\phi_R)}C_R} + \frac{b(\phi_R){C_R}^m}{S + b(\phi_R)C_R + \omega_{D(\phi_L)}C_L}, \tag{6.5}$$

where ϕ_L and ϕ_R refer to the orientation difference (in degrees) between the mechanism's peak sensitivity and the stimulus in the left and right eye respectively. All other terms are as described above, and in Chapter 3, with stage two of the model unchanged.

As previously, four phase sensitive mechanisms are considered (see section 6.1.4), all of which are optimally tuned to the test orientation (90°). The antiphase condition is handled in the same manner as for Experiment 11 (section 6.1.4). The best fit of this model is shown in Figure 6.8, with parameters given in Table 6.1. Model parameters not given in the table (*p, q, Z* and *m*) were fixed at the values from Meese *et al.* (2006), as previously.

For both subjects, the model fit is very good, describing well the fall-off in threshold elevation as mask and test orientations diverge, and producing acceptable RMS errors (average = 1.53dB). As found empirically, the two phase conditions are equivalent over most of the range of mask orientations, differing only when mask and test are very similar in orientation. Here, the model does capture the greater masking in the in-phase condition. However, rather than a single peak when mask and test have the same orientation, there are two peaks, a few degrees above and below the midpoint of the function.

This unexpected behaviour occurs because of complex interactions between the mask and test channels across the two intervals of the experiment. To gain insight into this,

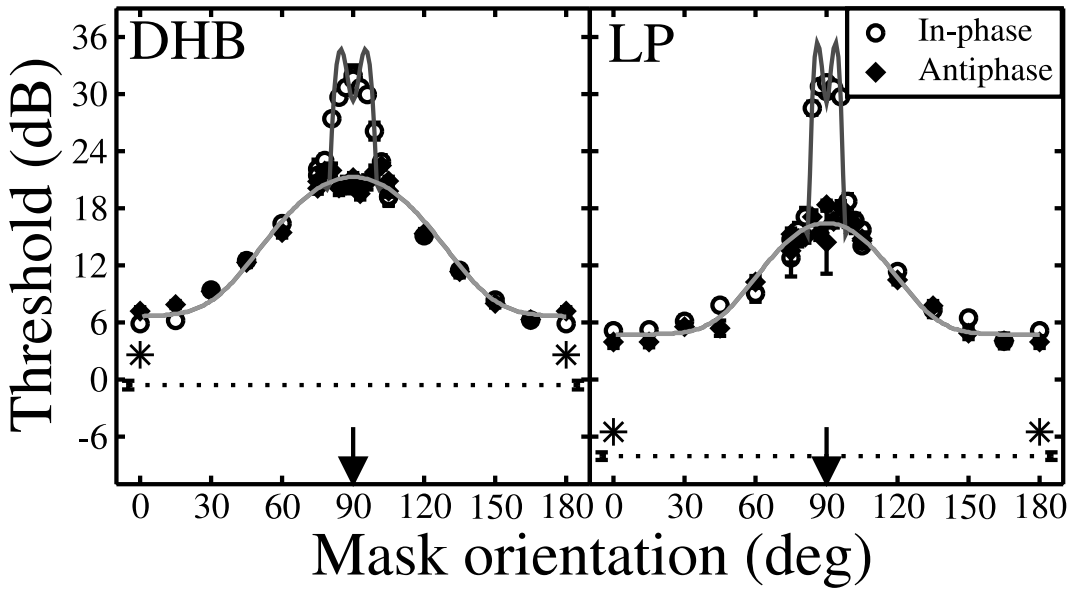

Figure 6.8: Best fits of the orientation model. Data points are replotted from Figure 6.6 (dotted lines are the empirical baseline thresholds). Curves and asterisks are the model predictions for the masking data and the baseline threshold, respectively. Parameters and RMS errors are given in Table 6.1.

it is instructive to look at activity at different points in the model, for different stimulus presentations. Figure 6.9 shows model activation in test (upper row) and null (middle row) intervals as mask orientation is varied and mask and test contrasts are kept constant at 30dB. In the top row, the leftmost panel shows activation in the test channel just after stage one. This is substantial at distant orientations (when masking is weak), but decreases as mask orientation approaches 90° according to the broad tuning of the dichoptic gain term, $\omega_{D(\phi)}$. The mask eye of the detecting mechanism (middle column) shows no activation at distant orientations, and some weak activation close to 90° determined by the narrow tuning of the sensitivity parameter, $b(\phi)$. The rightmost column shows the binocular sum of mask and test channels, prior to further processing at stage two of the model.

The second row of Figure 6.9 shows model activation in the null interval. The test eye is presented with mean luminance, so activation is zero (left panel). The mask eye shows narrowly tuned activation (defined by $b(\phi)$). This is stronger than the equivalent mask channel activation in the test interval because there is no dichoptic suppression from the test channel. The rightmost column shows binocular activity for the null interval.

The lower row of Figure 6.9 shows the activity at the output stage of the model (after stage two) in both null and test intervals (left and centre panels). Stage two is a further normalization stage, which amplifies activation, but does not change the shape of the functions. The lower right panel shows the difference between test and null intervals, which gives the final model response according to equation 3.2. This resembles an inverted version of the

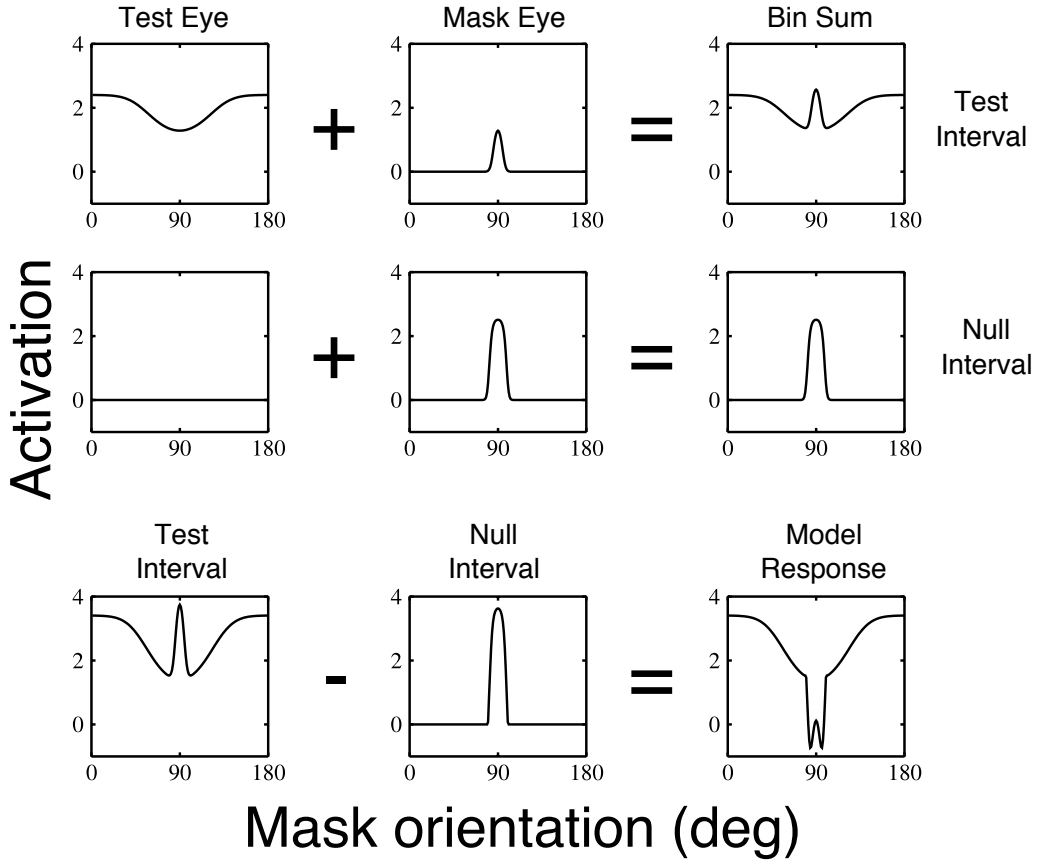

Figure 6.9: Activation in the orientation model, in both intervals of the experiment and at various locations. Both mask and test had contrasts of 30dB, with the test orientation fixed at 90° and the mask orientation given by the abscissae. The upper two rows show activation in the test and null intervals respectively, just after stage one in the test eye (left column) and mask eye (middle column), and after binocular summation (right column). The bottom row shows activation at the output stage in the test and null intervals, and their difference, which gives the final model response.

model threshold predictions shown in Figure 6.8, including the two peaks adjacent to the test orientation.

Inspection of Figure 6.9 reveals that the peaks arise because of the differencing operation used to calculate thresholds from the model, and the effect of nonlinear operations (gain controls) on the gaussian tuning functions. The absence of such peaks in the empirical data suggest either that the model is incorrect in some detail of its organisation or tuning, or that the decision process is more complex than the differencing operation used in modelling.

There is some indication from the antiphase condition that the latter may be the case, as observers are apparently able to ignore the activation in the mask channel, basing their decision purely on activation in the test channel (see section 6.1.4). It seems plausible that observers might use some sort of perceptual template-based decision strategy, such as responding to the interval which contains two stimuli (an AND operator), when mask and test are similar in orientation. However, implementing this computationally would require a departure from the traditional method of modelling contrast detection and discrimination data (i.e. Legge and Foley, 1980; Foley, 1994) which is consistent with the established tenets of signal detection theory (Green and Swets, 1966). Such possibilities are not pursued further here.

Another avenue which appeared promising was to introduce additional mechanisms sensitive to different orientations. These mechanisms were often more responsive than the optimal mechanism at orientations close to that of the test (consider small lateral shifts of the plot in the lower right panel of Figure 6.9). This improved the fit without introducing additional free parameters, by reducing the peaks (using an off-channel looking process). However, it did not remove them entirely, and so is not shown.

Despite its shortcomings, the model is still informative as it characterises the gain pool and also provides bandwidth estimates, as discussed in the following section.

Bandwidth estimates

The model parameters σ and ψ can give a good objective estimate of the bandwidths of interocular suppression and binocular summation. The bandwidth (full-width-at-half-height) of a gaussian is 2.35 times its standard deviation (Graham, 1989). This yields bandwidth estimates between 45 and 54° for interocular suppression, and between 5 and 8° for binocular summation, depending on subject. These are discussed further in section 6.6.

6.4.2 Spatial Frequency model

The model developed above can also be adapted to the spatial frequency domain. The spatial frequency masking data are plotted on a logarithmic axis, but they are well approximated by a gaussian which is a linear function of spatial frequency.

An additional parameter is necessary to shift the peak of the antiphase masking function slightly to the right of the test frequency. However, it is still possible to anchor the model such that when mask and test frequencies are equal, the suppressive weight is unity, consistent with the basic form of the model. Thus, the gaussian function of spatial frequency becomes,

$$\omega_{D(\rho)} = \frac{w + e^{\frac{(\rho - |R|)^2}{2\sigma^2}}}{w + e^{\frac{(-|R|)^2}{2\sigma^2}}}, \tag{6.6}$$

where w governs vertical elevation, $|R|$ determines rightward shift, and ρ is the difference between the peak sensitivity of the mechanism and the stimulus spatial frequency (in cpd). As the functions for spatial frequency are not symmetrical about the midpoint, the sign, as well as the magnitude, is important when calculating ρ. It is defined as,

$$\rho = f_{PEAK} - f_{STIM}, \tag{6.7}$$

where f is either the spatial frequency of the stimulus ($_{STIM}$) or the optimal frequency of the detecting mechanism ($_{PEAK}$). The shift parameter ($|R|$) allows the maximum to be at a point other than when the frequencies are equal. Positive values of $|R|$ shift the function to the right, and negative values shift it to the left. Since the antiphase masking functions for both observers have a peak above the test spatial frequency, $|R|$ was constrained to take on only positive values. The remainder of the model is equivalent to that defined in equations 6.4 and 6.5, but recast for spatial frequency in the manner described above.

The spatial frequency model has six free parameters; $k, S, w, |R|, \sigma$ and ψ. Parameter values and RMS errors are given in Table 6.2, with the best fits shown in Figure 6.10. Again, the fit is good for the antiphase data[2], but suffers from the same additional peaks for in-phase presentation as the orientation model (see section 6.4.1). As with orientation, these peaks are unfortunate but unavoidable without substantial reworking of the decision stage of the model, which would most likely require several unfounded assumptions.

[2]The greater RMS error found for subject DHB is largely attributable to the difference between the two conditions at low mask spatial frequencies, which the model does not capture.

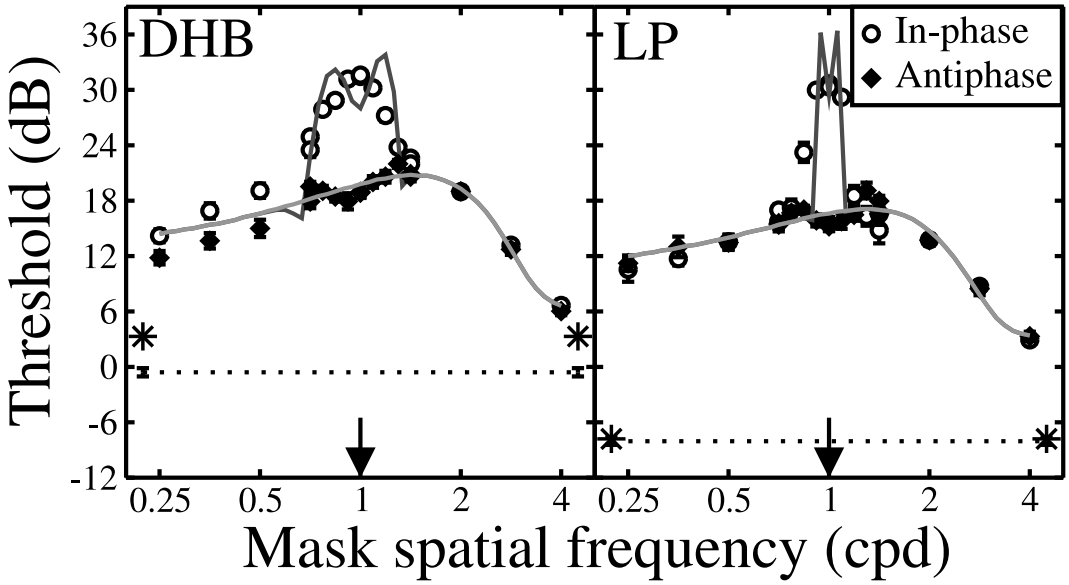

Figure 6.10: Best fits the spatial frequency model. Data points are replotted from Figure 6.5 (dotted lines are the empirical baseline thresholds). Curves and asterisks are the model predictions for the masking data and the baseline threshold, respectively. Parameters and RMS errors are given in Table 6.2.

Subject	Figure	σ	ψ	w	$\lvert R\rvert$	S	k	RMS error (dB)
DHB	6.10A	0.766	0.115	0.04	0.445	2.40	0.0129	1.98
LP	6.10B	0.757	0.031	0.11	0.299	0.66	0.0008	1.55

Table 6.2: Parameters and RMS errors of the spatial frequency model. Meanings of the parameter terms are discussed in the text.

Bandwidth estimates

Using the standard deviation parameters given in Table 6.2 (σ and ψ), bandwidth estimates can be calculated in cycles per degree. This is less straightforward than for orientation, as the spatial frequency functions are asymmetrical on logarithmic axes. For interocular suppression, the bandwidth is around 1.8cpd, which corresponds to approximately 1 octave above, and > 2 octaves below the test spatial frequency, or a total bandwidth of >2 octaves. The sensitivity parameter (ψ), which gives an estimate of the tuning of binocular summation, has a bandwidth of around 0.2cpd, which is <0.5 octaves.

6.5 Monocular tuning functions

Previous experiments have found substantial differences in the characteristics of monocular and dichoptic masking (see Chapter 5), which was taken as evidence that they have distinct neural substrates. Performing a monocular version of Experiments 12 and 13 would permit comparison with the dichoptic data.

Spatial frequency and orientation tuning has been measured previously for monocular and binocular masking. Legge (1979) performed a monocular version of his dichoptic spatial frequency tuning experiment, and concluded that the monocular functions were qualitatively similar, but more broadly tuned than their dichoptic counterparts. However, as with dichoptic masking, the functions were sparsely sampled.

Phillips and Wilson (1984) report the complementary experiment for orientation tuning over a range of spatial frequencies, also using monocular presentation. Their data, although finely sampled, omitted the condition where mask and test have equal orientation, and also only report thresholds up to an orientation difference of 45°. The Phillips and Wilson (1984) study was extended by Holmes (2003, see also Meese and Holmes 2003), who explored the full range of orientation differences between mask and test (0 - 90°) for a range of spatial and temporal frequencies. Under some conditions, orientation tuning was profound, with thresholds falling by 18dB (a factor of 8) over the 90° range of mask orientations. However, Holmes' data were gathered binocularly, and did not use a temporal envelope comparable to the 200ms presentation used here.

Clearly there is a need for monocular masking data which can be compared directly to the dichoptic results of Experiments 12 and 13. To address this, monocular tuning functions were measured for two subjects, using the same stimuli and methodology as for dichoptic presentation. As monocular presentation requires physical superimposition of stimuli, the antiphase manipulation is not viable. Thus, data were gathered only for in-

phase presentation.

6.5.1 Results

Monocular spatial frequency masking functions are shown for both subjects in Figure 6.11. Masking is substantial at all mask frequencies tested, raising detection thresholds by 6dB (a factor of 2) or more. For subject RJS, the function does not show strong tuning. For DHB, there is marked attenuation towards the more extreme spatial frequencies. Surprisingly, there is also a reduction in the level of masking when the mask and test frequencies are equal, as compared with the adjacent mask frequencies. This finding is not apparent for subject RJS, however, it is not without precedent in the literature. The 1cpd monocular spatial frequency tuning data of Legge (1979) also show a reduction in the level of masking at the test frequency, relative to those adjacent to it.

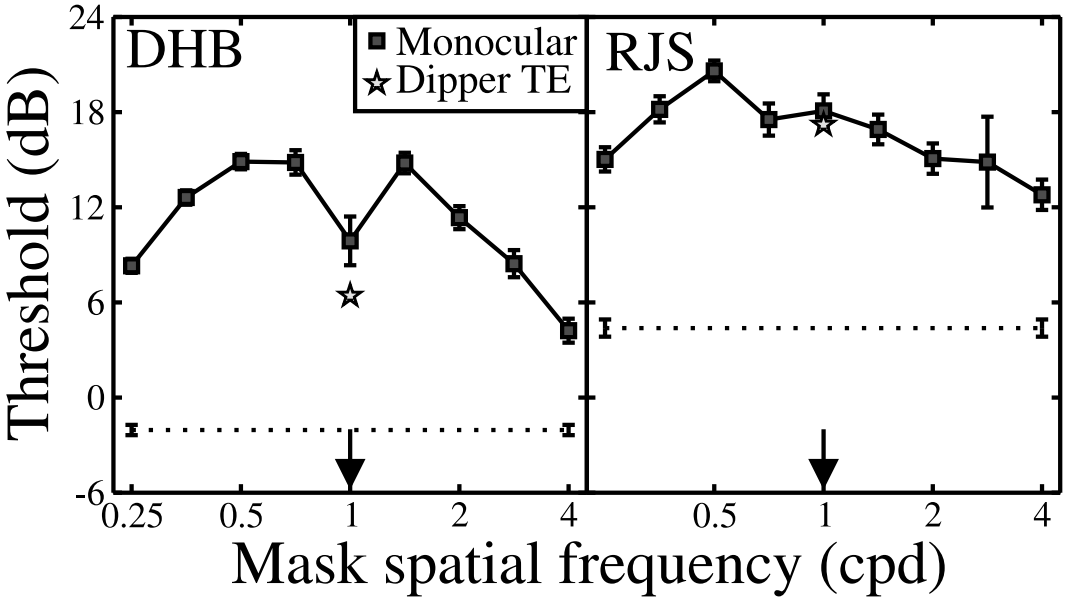

Figure 6.11: Monocular spatial frequency tuning functions for two observers. Filled squares are monocular thresholds in the presence of a horizontal 30dB mask, with mask frequency given on the abscissa. The test was always a horizontal grating at 1cpd. The horizontal dotted line gives detection threshold in the absence of a mask. The small star is the threshold elevation expected for each subject based on the results of Experiment 1 (Chapter 3). Error bars give the standard error of the probit fit.

Figure 6.12 shows the monocular orientation masking functions, which are similar in form to those reported elsewhere in the literature (i.e. Phillips and Wilson, 1984). Masking is strong (>12dB threshold elevation) at orientations close to that of the test, but shows strong tuning, reducing to ∼3dB of threshold elevation when mask and test are orthogonal. The function for DHB is somewhat broader than that for RJS. Again, subject DHB exhibits a curious reduction in masking when mask and test are co-oriented. There is some hint of

this behaviour for subject RJS, but it is nowhere near as profound.

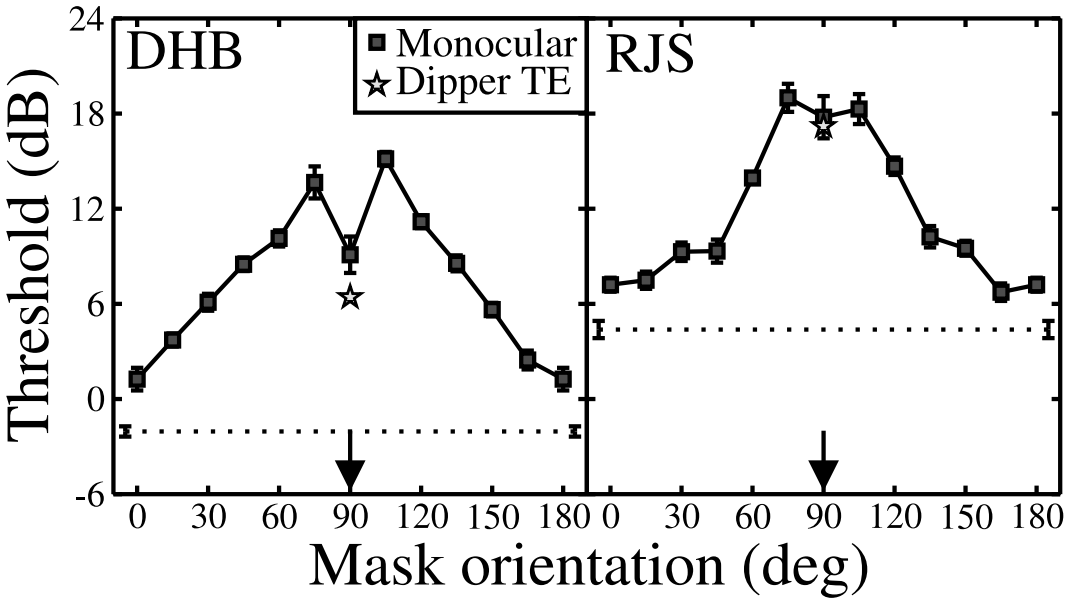

Figure 6.12: Monocular orientation tuning functions. Layout is the same as for Figure 6.11, except that mask orientation was varied.

It seems likely that subjects use different cues to make their responses when mask and test have the same properties compared with when they have very different properties. Since the stimuli sum together physically, the task is contrast discrimination when mask and test are the same. However, at more extreme mask orientations and spatial frequencies, the task is contrast detection. If one subject were not using the available cues optimally to perform the appropriate task, it might explain the discrepancy between the data of DHB and RJS at the midpoints of the functions.

To assess this, the appropriate data from the contrast discrimination (dipper) experiment (Experiment 1) of Chapter 3 are plotted in Figures 6.11 and 6.12 (stars)[3]. It is clear that both subjects are reasonably consistent across experiments, which took place more than two years apart. This indicates that the unexpected drop in the functions for DHB is not due to some change in strategy across experiments, nor to a poor estimate of threshold at this point.

It is not clear what causes this unusual effect. One possible explanation is that subject DHB is performing sub-optimally at mask orientations just above and below the orientation of the test. However, it is also possible that the tuning of the suppressive gain pool differs between subjects, resulting in different levels of masking for different mask types. As the data approximately resemble the additional peaks produced by the dichoptic model (see section 6.4.1), attempts were made to incorporate the data into a similar model. However,

[3]To calculate this value, the monocular threshold elevation at 30dB mask contrast in the dipper experiment was added to the detection threshold for the present experiment.

these were unsuccessful in accounting for the additional peaks in any plausible fashion, and their existence remains unexplained.

6.5.2 Comparison with dichoptic model

To compare the tuning of monocular and dichoptic suppression, the monocular data are plotted alongside the dichoptic model predictions in Figure 6.13. The model curves were generated for the dichoptic antiphase condition, using the parameters for DHB given in Tables 6.1 and 6.2, and provide a summary of the dichoptic results. For ease of comparison, both model and data are normalized to their appropriate baseline detection thresholds, and expressed as threshold elevation.

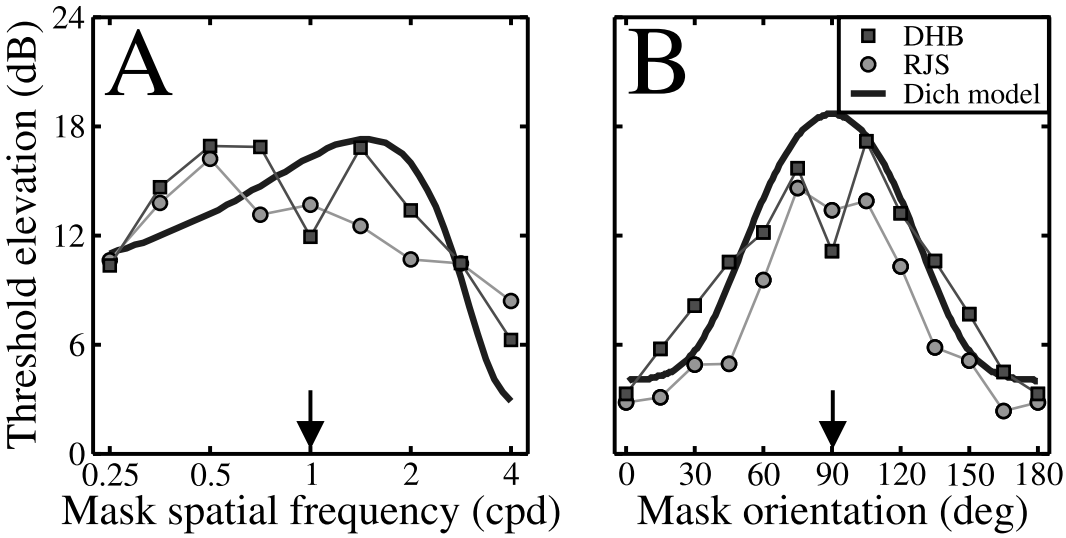

Figure 6.13: Comparison of monocular tuning data and the dichoptic model. Data are replotted as threshold elevation for DHB (dark squares) and RJS (light circles) for spatial frequency (A) and orientation (B) tuning. The curve is the prediction of the dichoptic model, developed in section 6.4, for the antiphase dichoptic condition.

The monocular spatial frequency functions shown in Figure 6.13A are more broadly tuned than their dichoptic counterpart, consistent with the findings of Legge (1979). The falloff in threshold elevation is noticeably shallower at the higher spatial frequencies for monocular masking. Threshold elevation is greater at 4cpd, and also at the lower spatial frequencies in between 0.25 and 1cpd.

For orientation tuning (Figure 6.13B), the monocular data are surprisingly similar to the dichoptic model, which falls between the data of the two subjects over much of the range. This is surprising given the substantial differences found between monocular and dichoptic masking elsewhere in this thesis (i.e. Chapters 4 and 5) and in single-cell studies (Li *et al.*, 2005; Sengpiel and Vorobyov, 2005).

The main difference between monocular and dichoptic functions is when mask and test are equal in orientation and spatial frequency. Here, monocular masking is substantially less severe than dichoptic masking, by up to 6dB. As discussed above, the cause of this threshold difference in the context of the entire tuning function is unknown.

6.6 Chapter discussion

Three experiments were carried out to explore the phase, orientation, and spatial frequency tuning of dichoptic masking. In Experiment 11, presenting stimuli in dichoptic antiphase was found to eliminate facilitation at low mask contrasts, and produce weaker masking at high mask contrasts, compared with in-phase presentation. The model indicates that antiphase presentation measures interocular suppression directly, without any additional masking due to binocular summation. Experiments 12 and 13 measured spatial frequency and orientation tuning functions for dichoptic masking, both in-phase and in antiphase, so that bandwidths for interocular suppression and binocular summation could be inferred. Using this method, interocular suppression was found to be broadly tuned ($\sim$50°; >2 octaves), and binocular summation was narrowly tuned ($\sim$8°; <0.5 octaves). Tuning functions for monocular masking are also reported, which have similar orientation bandwidths, but broader spatial frequency bandwidths than for dichoptic masking.

From the dichoptic data, binocular summation appears to be very narrowly tuned indeed. However, the use of large field gratings (5°) limits the conclusions which can be drawn from this finding. For phase-sensitive processes such as binocular combination, probability summation over area can artificially narrow the tuning of any summation effects (Bergen *et al.*, 1979). To use the present orientation tuning paradigm as an example, imagine that when mask and test are both horizontal the observer integrates information over the whole stimulus area. Binocular summation occurs across the whole stimulus in this condition, which in the masking paradigm produces strong threshold elevation. When the mask is rotated by 15°, there is a relatively small change in the overlap between the gratings in the centre of the stimulus, so binocular summation would remain strong if only this region were monitored. However, at the extremes of the grating, the phase relationship between mask and test changes, and large portions may become out of phase. This reduces the apparent level of binocular summation, and may result in narrower estimates of tuning than are possessed by the underlying physiology.

This effect, which applies equally in the spatial frequency domain, prevents any strong conclusions from being drawn about the neuronal bandwidth of binocular summation. Future work aimed at addressing the bandwidths of binocular summation should use smaller

stimuli to avoid such problems with area summation. However, as interocular suppression is apparently phase insensitive it is presumably immune to such effects, so the broad bandwidths estimated for this process should be robust for different stimulus areas.

Despite this limitation, there are some strong similarities between the bandwidths for binocular summation measured here, and those from related studies in the literature. Blakemore (1970) presented vertical gratings (3°x2.25°, 0.5-15cpd) to the two eyes which differed in their spatial frequency (see Figure 6.14A). The most striking observation was that subjects reported a strong depth percept, which was dependent on the interocular spatial frequency ratio. When this occurred, subjects saw a grating of a single spatial frequency, and did not experience binocular rivalry. This fusion took place over a limited range of spatial frequency ratios, from 0.7 to 1.4. In other words, when the gratings differed in spatial frequency by a factor of $<\sqrt{2}$, or about half an octave, they were binocularly fused into a single image. This agrees well with the bandwidth of binocular summation reported for Experiment 12.

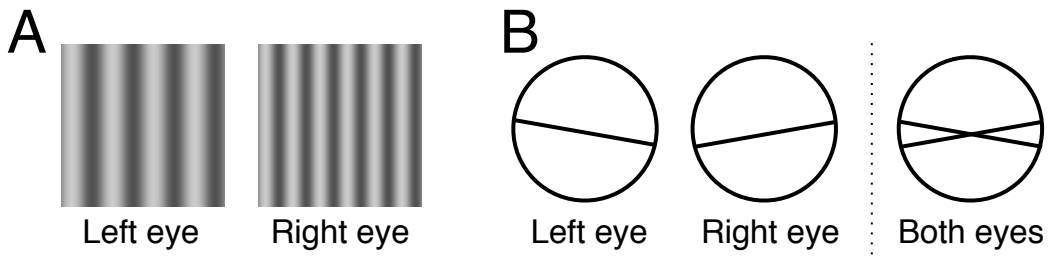

Figure 6.14: Illustration of stimuli used to measure binocular fusion. A) shows gratings differing in spatial frequency, presented to left and right eyes. Blakemore (1970) found that when spatial frequency differed by a factor of $\sqrt{2}$ or less, subjects saw a single grating sloping in depth. B) shows stimuli from an analagous experiment performed for orientation by Kertesz and Jones (1970). When the orientation difference was $<10^{\circ}$, subjects did not see two lines (as depicted), but reported a single horizontal line.

At around the same time, Kertesz and Jones (1970) set out to measure the range of interocular orientation differences over which subjects could make torsional (rotational) eye movements in order to fuse two line stimuli (10.5° of arc in length) differing in orientation (see Figure 6.14B). Subjects reported strong fusion over a 10° range of orientations, but surprisingly made no rotational eye movements (as measured by an accurate infra-red sensor). They conclude that some central process is responsible for the 'cyclofusion' experienced by their subjects. This small range in orientation difference corresponds well to the narrow orientation bandwidth of binocular summation found here.

6.6.1 Predictions of the dichoptic model

Although constrained by data where either mask orientation or spatial frequency was varied, the model can make predictions about any combination of mask properties. The dichoptic

weight at any mask orientation and spatial frequency is assumed to be the product of the two individual weights ($\omega_{D(\phi)} \times \omega_{D(\rho)}$). This assumption (polar separability) is intuitive and straightforward, but is by no means the only combination rule possible (for example, Phillips and Wilson (1984) used Cartesian separability to model linear filter properties of visual mechanisms).

The surface plots in Figure 6.15 summarise the behaviour of the full model. Panel A shows masking in the antiphase condition, characterised by a smooth continuous surface. Panel B shows the equivalent prediction for the in-phase case, where an additional raised 'nipple' is apparent at spatial frequencies and orientations close to that of the test. Panel C shows the same area of masking space but for a much lower mask contrast (5dB). Here, masks similar to the test cause facilitation, resulting in a dramatic reduction in thresholds, consistent with Figures 6.1 and 6.3 (see also Chapters 3 and 7). The accuracy of these model predictions at low mask contrasts remains to be tested experimentally.

The model can be used to predict the spatial frequency tuning of orthogonal masking, which was measured experimentally in Experiment 9 (Chapter 5). Figure 6.16 shows the data from that experiment, plotted as threshold elevation averaged across the two observers. The black line is a direct model prediction, using the parameters for DHB derived above. The model predicts less masking than is found empirically. However, as was noted in Chapter 5, the temporally modulated (4Hz) stimuli used in the orthogonal masking experiment appear to elicit more masking than do the static stimuli used in this chapter. This can be corrected for by normalizing the model prediction to the leftmost data point (grey function). The model then produces a reasonable approximation of the shape of the tuning function.

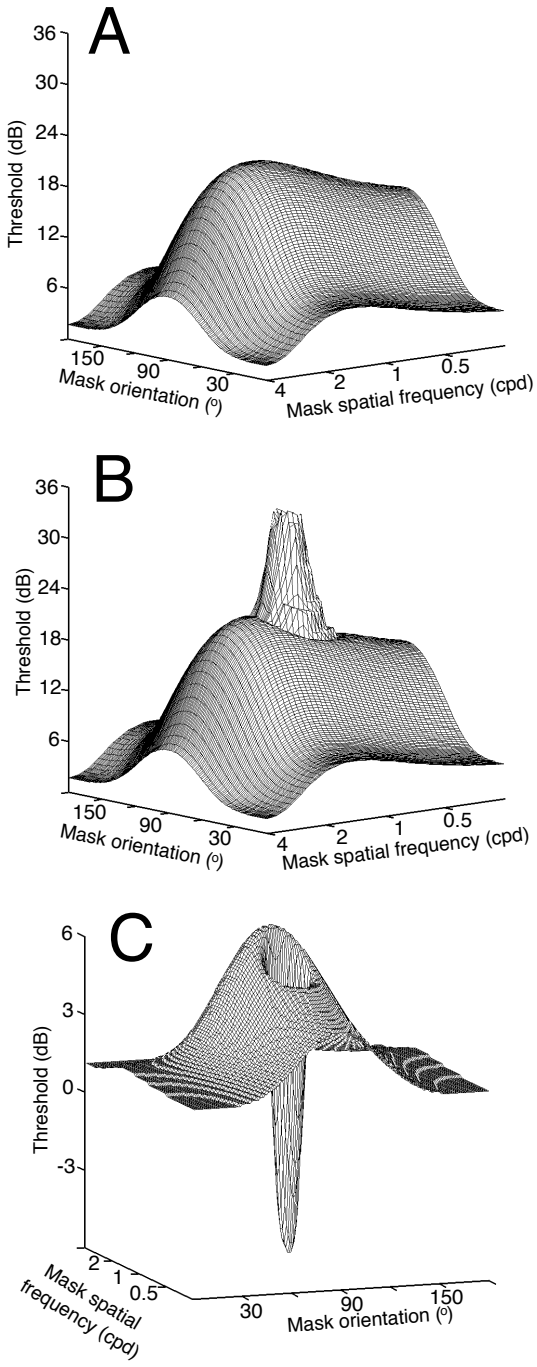

Figure 6.15: Dichoptic masking surfaces, generated using the model. All panels show the model predictions for dichoptic masking over a range of mask orientations and spatial frequencies. The test was always a 1cpd horizontal (90°) grating. In A) mask and test are 180° out of phase, and mask contrast is 30dB. In B) the conditions are the same except that the stimuli are in-phase. In C), stimuli are again in-phase, but the mask contrast is 5dB. There is a clear transition from masking to facilitation when mask and test have similar properties.

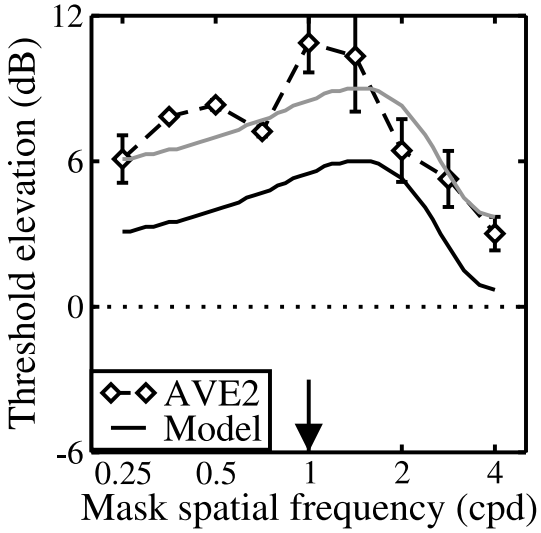

Figure 6.16: Predictions of the model for orthogonal masks (90°) of varying spatial frequencies (black curve). The symbols are data from the dichoptic condition of Experiment 9 (Chapter 5), replotted as elevation above detection threshold averaged across two subjects. Note that for this experiment all stimuli were temporally modulated at 4Hz, whereas the model is based on thresholds for static stimuli presented for 200ms. The grey curve is the model prediction normalized to the leftmost data point.

Chapter 7

Amblyopia

7.1 Amblyopic contrast vision

It has long been known that humans with amblyopia exhibit reduced contrast sensitivity in the affected eye (Hess, 1979; Bradley and Freeman, 1981; Hess, Bradley and Piotrowski, 1983; Hess, 1983; Asper, Crewther and Crewther, 2000). This loss could be due to several factors, including fewer active cells (Levi, Klein and Yap, 1987), inhibition between the two eyes (Sengpiel and Blakemore, 1996), or disorganisation of neurones (Hess and Field, 1994). The sensitivity difference between the eyes increases with spatial frequency. Some amblyopes show no difference at lower spatial scales, whilst others show sensitivity reductions at all spatial frequencies (Hess and Howell, 1977).

Binocular summation is also affected (Lema and Blake, 1977; Levi *et al.*, 1979, 1980). In many subjects, no summation is observed, and binocular contrast thresholds are determined by the normal eye (a summation ratio of $\sim$1). Some studies (i.e. Westendorf, Langston, Chambers and Allegretti, 1978) have found summation ratios of the magnitude expected from probability summation ($\sim$1.2; Pirenne, 1943; Eriksen, 1966; Tyler and Chen, 2000). Pardhan and Gilchrist (1992) report data from four subjects who show binocular summation ratios of less than one at high spatial frequencies. These data imply that binocular performance is substantially worse than performance with one eye only (see also Hood and Morrison, 2002; Holopigian *et al.*, 1986), suggesting the presence of some suppressive process occurring between the eyes. Levi *et al.* (1980) came to a similar conclusion after finding that dichoptic masking is still present in amblyopes, despite an absence of binocular summation.

Interestingly, there is some evidence that binocular summation ratios of less than one can also be observed in normal subjects, when the input to one eye is attenuated by a neutral density filter (Gilchrist and McIver, 1985; Heravian-Shandiz, Douthwaite and Jenkins, 1991).

It is possible that this procedure could mimic the behaviour of amblyopes in other ways also. Two studies report that dominance durations in a binocular rivalry paradigm for normal subjects wearing ND filters resemble those of amblyopic subjects (de Belsunce and Sireteanu, 1991; Leonards and Sireteanu, 1993). Interestingly, when amblyopes wore a filter over their good eye, their data became more like that of normals. This suggests that a reduction of luminance in one eye may be a viable model of amblyopia.

Harrad and Hess (1992) measured dichoptic masking in a group of amblyopes over a range of spatial frequencies. They hypothesised that the interocular suppression which causes dichoptic masking might be unaffected by amblyopia. If this were true, dichoptic masking functions for the two eyes should superimpose when plotted on threshold-normalized log axes. This was clearly not the case for the majority of their subjects. Instead, the slopes of the masking functions differed between the two eyes. For strabismic amblyopes, when the test was presented to the amblyopic eye, and the mask to the normal eye, masking was stronger than in normals. When mask and test eyes were reversed, masking was substantially weaker than in normals, and sometimes absent altogether. Harrad and Hess (1992) concluded that in strabismics, there is greater suppression of the amblyopic eye by the normal eye, and weaker suppression in the opposite direction, than is found in normals.

Meese and Hess (2004) proposed a model of binocular contrast processing with an early gain control, occurring before binocular summation, which allowed for interocular suppression of signals (see Chapter 3 for details). They suggest that this configuration could accommodate results like those of Harrad and Hess (1992) by adjusting the weights of interocular suppression between the eyes. In principal, this should adjust the steepness of the dichoptic masking function, making it steeper as the weights increased, and shallower as they decreased. However, this prediction has yet to be explored systematically, or tested on empirical data.

The present study (Experiment 14) is similar to that of Harrad and Hess (1992) in that it measures dichoptic masking in strabismic amblyopes. However, in order to obtain a clearer picture of amblyopic contrast response, full monocular and binocular contrast discrimination functions were also measured, as there is some suggestion that these might also be abnormal in amblyopes (Ciuffreda and Fisher, 1987; Kiper and Kiorpes, 1994), although other studies have shown discrimination to be normal (Hess *et al.*, 1983). The inclusion of a binocular condition allows the level of binocular summation at threshold to be measured.

7.2 Methods

7.2.1 Apparatus and Stimuli

Stimuli were patches of sinusoidal grating, spatially limited by a raised cosine envelope with a central plateau of 3° and a cosine half-period of 1°. Their spatial frequency was 3cpd (15 full cycles per patch). Stimuli at 3cpd were used, as many amblyopes do not have substantial deficits at the lower spatial frequencies (Hess and Howell, 1977) used previously (i.e. 1cpd). For one subject (EGF), detection thresholds at 3cpd were unmeasurable in the amblyopic eye. For this subject, 0.5cpd gratings were used, with the same spatial envelope. All stimuli were displayed on a Clinton Monoray monitor, using a VSG 2/5 and a PC running the Liberator software.

Most previous studies of this kind have used large vertical sinusoidal gratings as stimuli, presented through a stereoscope (i.e. Harrad and Hess, 1992). However, small vergence movements, or misalignments of the eyes, can cause vertical stimuli to slip out of phase, especially at high spatial frequencies (Green and Blake, 1981). Many stereoscopes also exacerbate this problem, as they can require subjects to actively fuse together the images from the two eyes. Both of these concerns are particularly relevant when working with strabismic amblyopes, in whom the eyes are already misaligned.

To avoid these potential confounds, ferro-electric shutter goggles were used, and each subject's strabismus was measured using a Lyle Major Amblyoscope (Clement Clarke Ltd., UK) and corrected using a prism (see Table 7.1 for prism strengths and Hood and Morrison, 2002, for a detailed discussion on the use of prisms). Furthermore, only horizontal gratings were used, at 3 or 0.5cpd, to prevent misalignment of stimuli.

7.2.2 Procedure

Subjects were seated in a darkened room, 114cm from the display. The goggles were worn on the head, attached using an elasticated strap, and prisms were fixed to the front of the goggles to correct the strabismus. The appropriate prism strength (given in Table 7.1) was assessed for each subject before the experiment began (see above), such that when viewed with the prism, a pair of nonius lines appeared collinear.

Observers indicated, using a mouse, in which of two intervals the highest contrast grating was displayed. Each interval was marked by a beep, and feedback was given after each trial indicating a correct or incorrect response. Ten pedestal contrasts were used; -10 to 30dB in steps of 5dB, and 0%. Data were gathered in blocks of a given pedestal contrast. Within each block, five staircases were interleaved, measuring right and left monocular thresholds,

right and left dichoptic thresholds, and a threshold for binocular presentation. Each block took around ten minutes, and subjects were given the opportunity to rest between blocks.

Subjects repeated the experiment four times, apart from DHB (six repetitions) and EGF (three repetitions). Data were collapsed across session, but analysed separately for each eye, using probit analysis (Finney, 1971). DHB also completed the experiment with a 1.5 log unit neutral density filter in front of the left eye. This dramatically reduced the mean luminance (to around $1 cd/m^2$), and was intended to impair contrast sensitivity in the filtered eye.

Pooled averaging

Data from amblyopic observers is typically noisy and heterogeneous. In order to produce a set of data which summarises the overall trends, and is suitable for modelling, the results for the eight amblyopes were averaged using a pooling method similar to that of Burton (1981). It was first necessary to normalize both axes of the masking function to the appropriate detection thresholds. Monocular and binocular functions were normalized to their own detection thresholds, for both pedestal and test contrasts. For the dichoptic functions, the test contrasts were normalized to detection threshold for the tested eye, and the pedestal contrasts were normalized to detection threshold of the *other* eye.

Because observers had different sensitivities, normalized pedestal contrasts were different for each observer, so thresholds were pooled across a limited contrast range (bin) before averaging (Burton, 1981). 5dB was chosen as the bin size, as this is equal to the pedestal contrast step size in the experiment. Substantially larger or smaller bin sizes produced results which did not resemble the data from individual subjects. This was due to either pooling across too wide a range of pedestal contrasts, or pooling across such a small range as to have very few samples.

After pooling, the data were 'de-normalized' to allow meaningful comparison with the raw and control data. This was achieved by using the binocular detection threshold as a baseline, and calculating the average threshold elevation relative to this, across subjects. This produced a threshold elevation for the amblyopic eye of around 12dB.

7.2.3 Observers

Eight strabismic amblyopes served as subjects. Their clinical and demographic details are given in Table 7.1. Normal optical correction was worn, and all amblyopes were psychophysically experienced, but naïve to the purposes of the experiment. All subjects were financially compensated for taking part, and were free to terminate the experiment at any time. Procedures adhered to the ethical guidelines of McGill University, where these experiments were

Observer	Age/Gender	Amblyopia	Prism	Refraction	Acuity	History
ADS	21/F	Right ET	10D	-0.5 DS	20/20	Detected age 4, patching age 4
		Strabismic		Ø	20/125	Surgery age 7
AR	47/M	Left ET	None	Ø	20/20	Detected age 20
		Strabismic		Ø	20/50	No treatment
EGF	56/M	Left ET	1D	+2.00+1.00 180°	20/32	Detected age 6
		Strabismic		+2.00+1.00 130°	20/250	Patching for 1-2 years
EMD	43/F	Left ET	3D	+0.75 DS	20/16	Detected age 6
		Strabismic		+0.75 DS	20/63	Patching for 1 year
JL	29/M	Left XT	20D	Ø	20/20	Detected age 4
		Mixed	(10D each eye)	+2.5 DS	20/40	No patching, no surgery
KDJ	22/M	RE XT	3D	+1 DS	20/50	Detected age 5, patching
		Strabismic		Ø	20/25	No surgery
ML	24/F	Right ET	3D	+1.0-0.75 90°	20/80	Detected age 5
		Mixed		-3.25 DS	20/25	Patching for 2 years
SH	24/F	Left XT	6D	-0.5 90°	20/32	Detected at birth
		Mixed		+2.5+2 180°	20/63	No patching, no surgery

Table 7.1: Demographic and clinical details of amblyopic subjects. Terminology: ET - esotropia; XT - exotropia; Ø - no refraction necessary. Acuity was measured using a standard LogMAR chart.

carried out.

The author (DHB, male, 24) served as a normal control subject at both spatial frequencies used, and repeated the experiment with a neutral density filter in front of the left eye (see above). Two additional normal subjects (RFH and BM) carried out only the detection threshold conditions, and the additional summation experiment (see section 7.4). Both were highly experienced psychophysically.

7.3 Results

Results for the control subject (DHB) are shown in Figure 7.1 - monocular data are plotted in the upper panels (A-C), dichoptic data are in the lower panels (D-F) and binocular data are plotted in all panels to provide a reference. At both spatial frequencies (panels A, B, D and E), the results replicate those discussed elsewhere (see Chapter 3, and Meese *et al.*, 2006). For all functions, facilitation is observed at low pedestal contrasts, and masking at higher pedestal contrasts. Substantial binocular summation is apparent at detection threshold (see figure caption). At 0.5cpd (panel A), the differences between left and right eyes are virtually non-existent. However, at 3cpd (panel B), the dominant right eye is around 1dB more sensitive than the left eye. This produces slight differences in the masking functions for the two eyes. Nevertheless, the results are well described by the Two Stage model (curves in Figure 7.1A, B, D & E), with only two free parameters, as described in the figure caption.

In Figure 7.1C and F, functions are shown for the same subject, but with a 1.5 log unit ND filter in front of the left eye. The filter substantially reduces the luminance to this eye (by a factor of 32), which increases detection thresholds by around 12dB. Monocular thresholds in the right eye are unaffected, however the binocular detection threshold increases substantially, almost up to the level of the right eye. This greatly reduces the binocular summation ratio, from 1.54 (no filter) to 1.16 (filter). The dipper functions are also affected. Dichoptic facilitation is no longer apparent in the right eye's function, and masking is substantially weaker (shifted rightwards). In the left eye, dichoptic masking becomes stronger, and dichoptic facilitation remains. The monocular dipper function for the left eye also changes, showing weaker facilitation with the filter than without it (compare panels B and C).

The results for all eight amblyopes are shown in Figures 7.2 (monocular and binocular) and 7.3 (dichoptic and binocular). An averaged data set was created, using the pooling method described in section 7.2.2, and is shown in the lower right panel (I) of Figures 7.2 and 7.3. The averaged data strongly resemble the data of DHB with an ND filter, shown in Figure 7.1C&F. These similarities are returned to in section 7.5. The pattern of monocular data for the amblyopes are analysed first, followed by the dichoptic masking

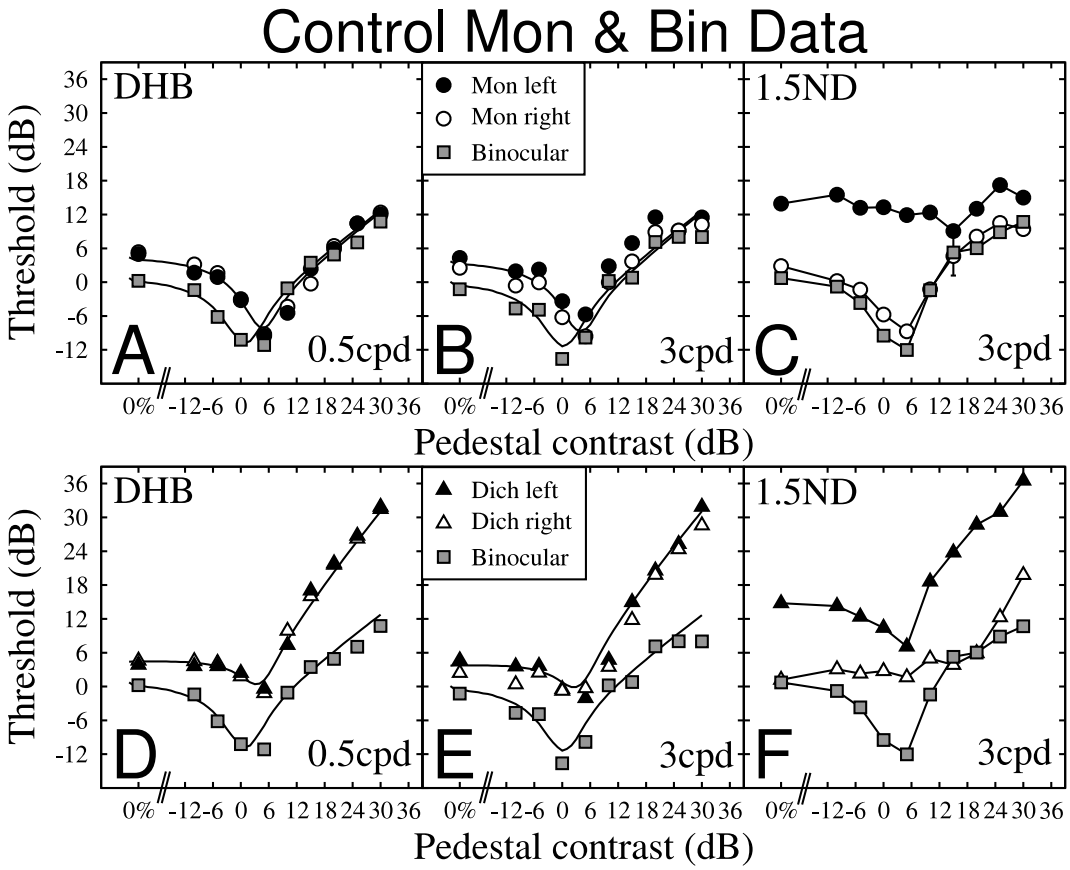

Figure 7.1: Contrast discrimination functions for normal control subject. Results are shown at 0.5cpd (A,D) and 3cpd (B,E), and also with a 1.5 log unit neutral density filter over the left eye (C,F). Upper panels (A-C) contain monocular and binocular data, and lower panels (D-F) contain dichoptic and (the same) binocular data. To minimise clutter, error bars (probit standard errors) are plotted only when they exceed 3dB. Binocular summation ratios, calculated by dividing the lowest monocular threshold by the binocular threshold, were 1.62 (A), 1.54 (B) and 1.16 (C). The curves in A, B, D and E are the best fits of the Two Stage model, using the parameters from Meese *et al.* (2006), with only k and S allowed to vary (0.5cpd: $k = 0.21$; $S = 1.36$; RMSe = 1.31dB. 3cpd: $k = 0.21$; $S = 1.20$; RMSe = 2.38dB).

data. A discussion of binocular summation at threshold is presented in the following section (7.4).

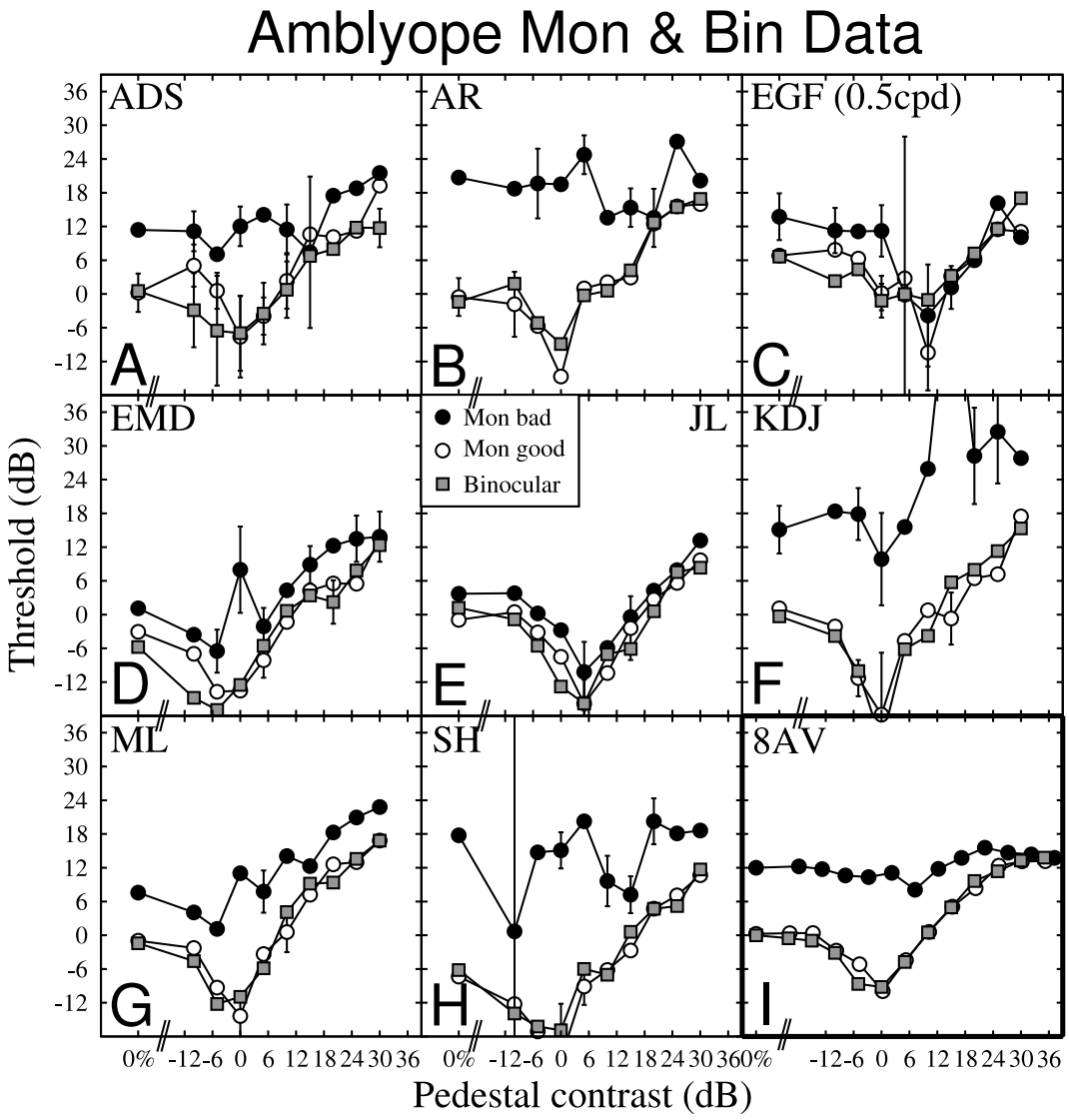

Figure 7.2: Monocular and binocular contrast discrimination functions for amblyopic observers. Panels are ordered alphabetically by first initial. Panel I contains data averaged across all eight observers, using the pooling method described in Section 7.2.2. Again, error bars are plotted only when they exceed 3dB, and give the standard error of the probit fit for individual subjects, or the standard error of the mean for the averaged data.

7.3.1 Monocular data

All subjects showed detection thresholds and monocular discrimination functions for their good eye which were comparable to those of the control subject. In the bad eye, however, detection thresholds were raised by between 4 and 23dB (factors of 1.6 and 14 respectively). This was accompanied by an elevation of the monocular dipper function for the bad eye, and generally a rightward shift of the dipper region. However, the dip was shallower than

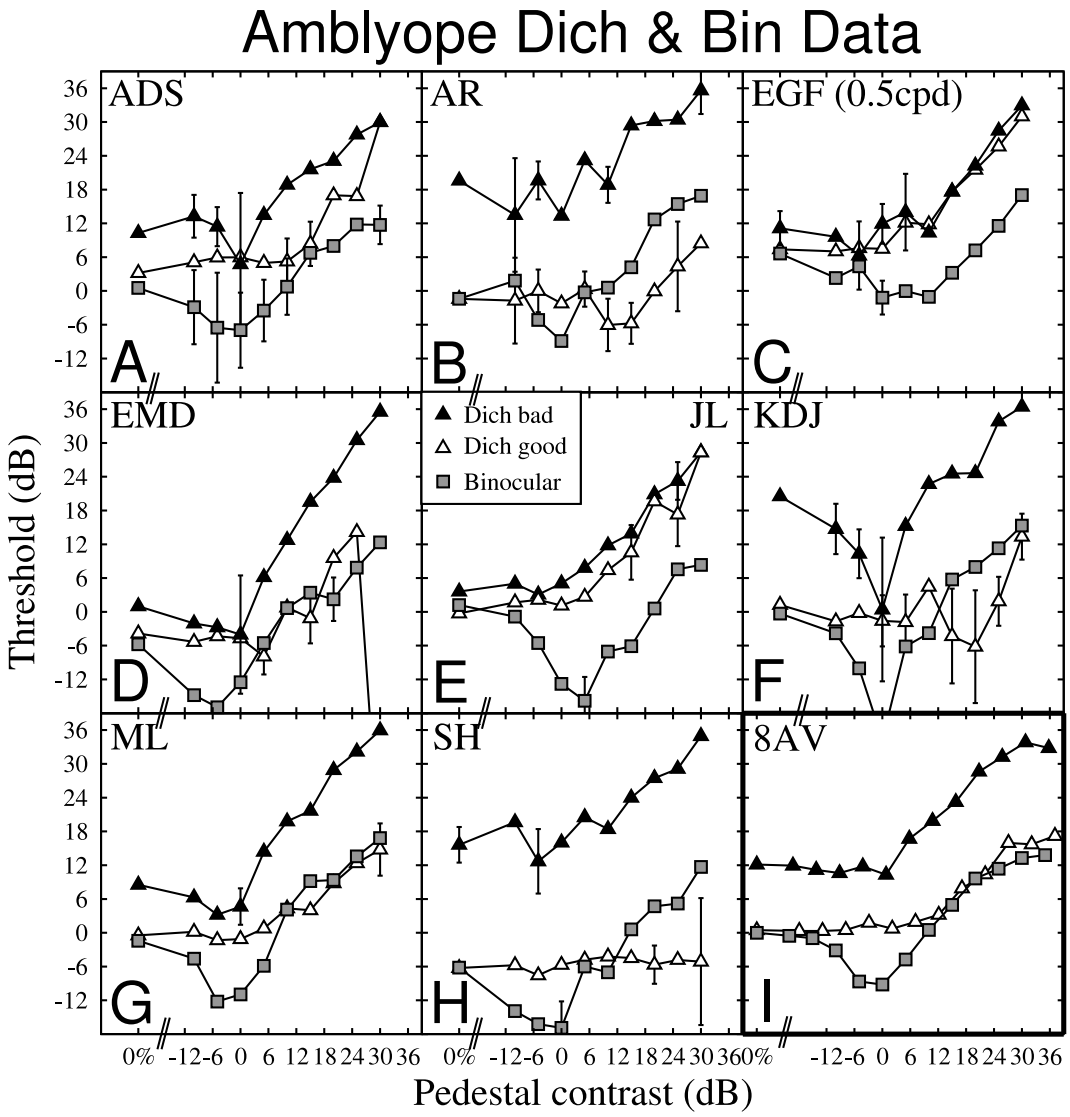

Figure 7.3: Dichoptic and binocular contrast discrimination functions for amblyopic observers. Details are the same as for Figure 7.2.

in the good eye, and in the averaged data it is greatly reduced (see Figure 7.4).

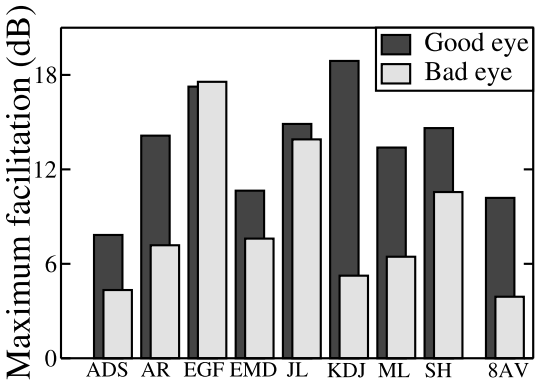

Figure 7.4: Maximal monocular facilitation in the good and bad eye for each amblyope, and for the averaged data. The amount of facilitation was calculated by subtracting the lowest threshold in the discrimination function from the baseline detection threshold for that eye (all values in dB). The only exception was for subject SH, where the lowest threshold was at -10dB pedestal contrast, and had an abnormally large error. This point was ignored, and the procedure carried out for the remainder of the function. Dark bars show facilitation in the good eye, and light bars show for the bad eye. In general, facilitation was reduced in the bad eye.

7.3.2 Dichoptic data

Dichoptic masking was evident in all subjects, and was substantial for most. Plotting masking data on normalized axes (see Figure 7.5) reveals that they mostly conform to the steep slope of around 1, found for normal observers (Legge, 1979). Where data points deviate from this, they tend to fall to the right of the oblique line, indicating weaker masking. This occurs mostly in the bad eye. (Slopes of the dichoptic functions are given in Table 7.2, calculated over the top 3 mask contrasts.) This is surprising, as Harrad and Hess (1992) found substantial deviations from the norm for many of their observers, with both steeper and shallower functions being common. For their strabismic subjects, dichoptic slopes tended to be steeper for the amblyopic eye - the opposite pattern to that found here. In many subjects, dichoptic facilitation was also found, particularly in the amblyopic eye (see points below the horizontal dotted line in Figure 7.5). This is of particular interest, as it is unlikely that dichoptic facilitation could occur in the absence of any excitatory binocular interaction (see Chapter 6).

7.4 Binocular summation at threshold

As noted above, binocular summation ratios for amblyopes are typically much lower than normal ($\geq\sqrt{2}$), having values much closer to one. Most studies (Pardhan and Gilchrist, 1992; Levi *et al.*, 1980; Vedamurthy, Suttle, Alexander and Asper, 2007; Pardhan and Whitaker,

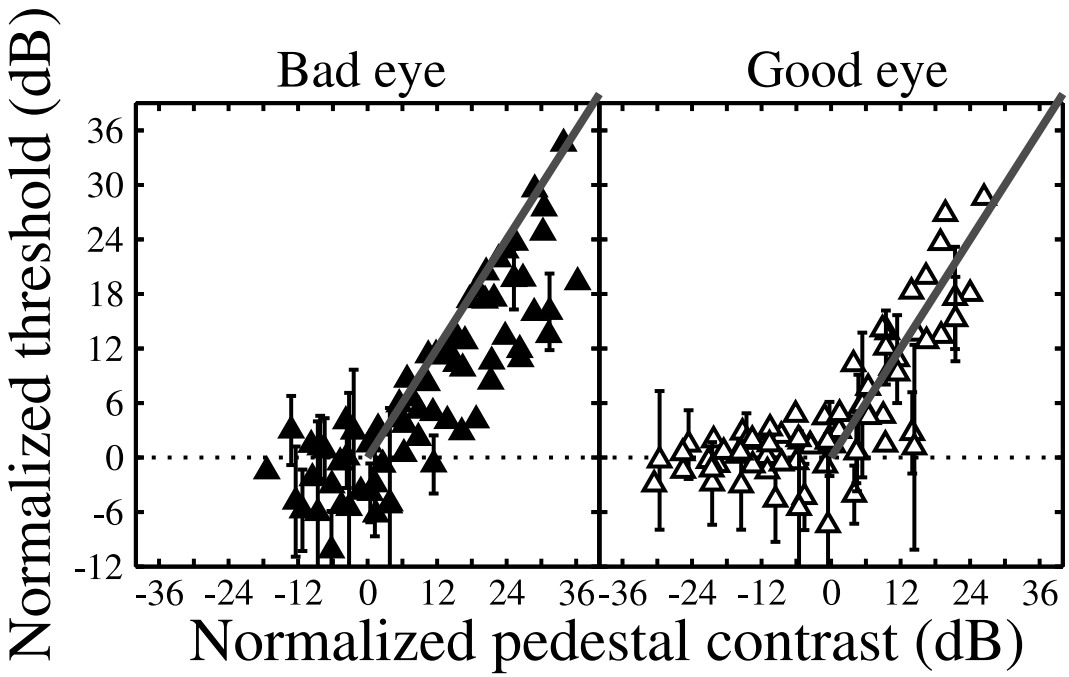

Figure 7.5: Normalized dichoptic masking data for all amblyopes. The oblique line of unity represents dichoptic masking in normal observers (Legge, 1979). The data generally follow this line, or sit to the right of it, indicating normal, or weaker-than-normal masking. Points below the horizontal dotted line are evidence of dichoptic facilitation. There is little or no evidence for unusually high levels of masking.

Observer	Bad eye slope	Good eye slope
DHB 3cpd	1.13	0.88
DHB 0.5cpd	1.03	0.96
DHB 3cpd ND	0.78	1.39
ADS	0.69	1.30
AR	0.55	0.85
EGF	1.07	0.95
EMD	1.17	1.53
JL	0.75	0.87
KDJ	1.18	1.95
ML	0.70	0.60
SH	0.75	0.06
Average	**0.86**	**1.01**

Table 7.2: Dichoptic masking slopes for all data sets. The dichoptic slopes were calculated by performing linear regression on the highest three mask contrasts in each dichoptic function (using dB units). Average values are for the amblyopes only (N=8).

2000; Hood and Morrison, 2002) have calculated the binocular summation ratio by comparing the binocular threshold to that of the better eye. The levels of binocular summation for each subject can be determined using this method[1], as shown in Figures 7.6 (controls) and 7.7 (amblyopes). It is clear that all amblyopes show reduced binocular summation, relative to the control subjects (Geometric means: 1.64 (controls), 1.07 (amblyopes)).

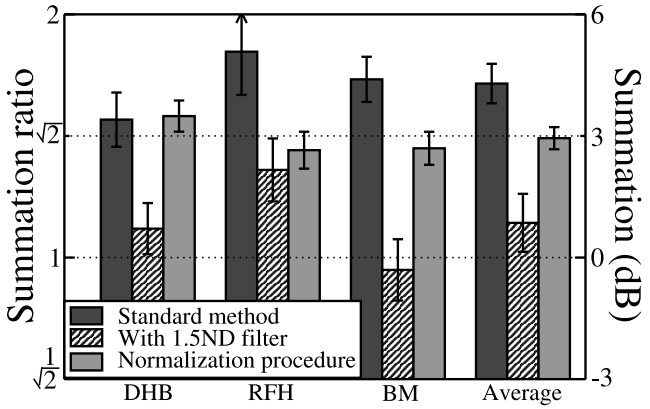

Figure 7.6: Binocular summation ratios for three normal observers. Dark bars are calculated using the standard method - dividing binocular sensitivity by the best monocular sensitivity. Striped bars were calculated using the same method, but when a 1.5 log unit ND filter was placed in front of the non-dominant eye. Grey bars are the results of the normalization procedure, also gathered using the filter. All ratios are the geometric mean of six estimates (sessions). The average is the geometric mean across all three observers. Error bars give the standard errors of the means.

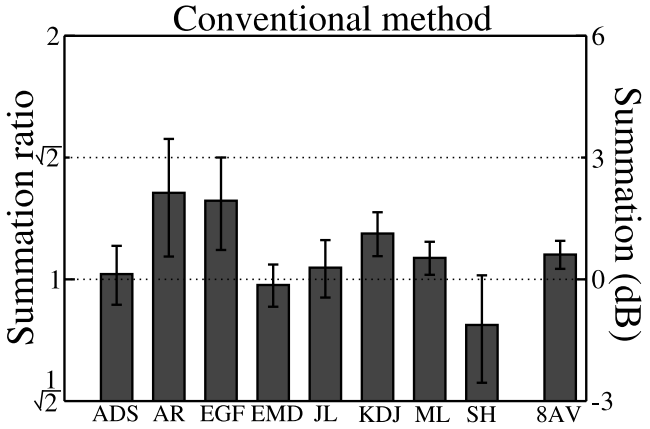

Figure 7.7: Binocular summation ratios for eight amblyopic observers, calculated using the standard method. Binocular sensitivity was divided by the best monocular sensitivity, on a session-by-session basis using the data from the main experiment (3 (EGF) or 4 (others) sessions), and the geometric mean calculated. The far right bar is the geometric mean across all eight observers. Error bars give the standard error of the mean. The spatial frequency was 3cpd for all subjects, except for EGF (0.5cpd).

However, this result does not necessarily mean that the amblyopes lack binocular cells, as has often been suggested, based on both human (Sireteanu, Fronius and Singer, 1981)

[1]Only the data from the monocular condition of the main experiment were used here, so that both the monocular and binocular thresholds were estimated from an equal number of staircases. Using the dichoptic data (which constitute the same condition in the absence of a pedestal) produced equivalent results (amblyope mean ratio 1.06).

and animal (Hubel and Wiesel, 1965; Sengpiel and Blakemore, 1996) studies. It is also possible that binocular summation is not observed due to the substantial sensitivity difference between the eyes. This is illustrated by considering the control subjects' summation ratios gathered with an ND filter over their non-dominant eye (Figure 7.6, striped bars). Here, binocular summation is reduced also (mean = 1.1), yet the mechanisms of binocular contrast integration are presumably intact in these normal subjects. Thus, the reduction in summation, to similar levels observed in the amblyopes, must be due to the sensitivity difference caused by the filter.

This suggests that simply comparing thresholds for the good eye and both eyes is an inappropriate method of determining the presence or absence of neural binocular summation. Instead, a psychophysical paradigm can be used which is designed for investigating summation of two stimuli for which observers differ in their sensitivity. In the summation paradigm, detection thresholds are first determined for each component (here, each eye). Then, a composite (here, binocular) stimulus is produced, such that the two components are equal in magnitude (in log units) relative to their own detection thresholds, and performance measured for this stimulus. Applying this method to binocular summation allows the sensitivity loss in the bad eye to be compensated for. This will provide a level 'playing field' for assessing the presence or absence of neural binocular summation in amblyopia.

Six of the eight amblyopes were available for this follow up experiment. Equipment and methods were as for the main experiment. Three control subjects also completed the experiment with ND filters in place as before[2]. The results for the control subjects are shown in Figure 7.6 (light bars). It is clear that summation is greater compared to the standard procedure, back to around normal levels (mean = 1.41 = $\sqrt{2}$). This validates the method, and shows that normalization reveals the true level of binocular summation when the eyes have different sensitivites.

For the amblyopes, a similar pattern is observed (Figure 7.8). Binocular summation is greater than $\sqrt{2}$ for five out of the six observers, and closely approaches it for the sixth (JL). The mean summation ratio is now 1.53. An ANOVA, comparing ratios from Figures 7.7 and 7.8 for the appropriate observers (n = 6), showed this to be a highly significant difference ($F = 35.7$, $p<0.00014$).

This finding has important theoretical implications for the understanding of amblyopia. It appears, for these subjects at least, that the neural mechanisms which underpin binocular summation remain intact. This means that amblyopes must still have a substantial number of functioning binocular cells, contrary to the findings of physiological studies on artificially strabismic kittens (Hubel and Wiesel, 1965). It is possible that the surgically induced

[2]As sensitivities in the two eyes of normal subjects are equivalent, performing this procedure without a filter produced equivalent results to the standard method.

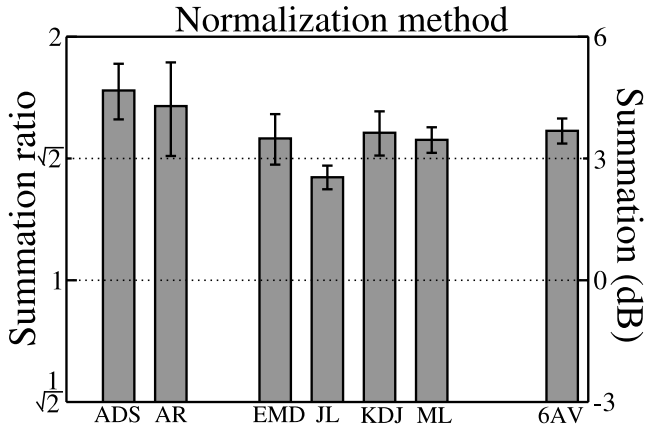

Figure 7.8: Binocular summation ratios for six amblyopic observers, calculated using the normalization method described in the text. Again, values are the geometric means across six sessions (or six observers for the average), with error bars giving the standard error. Gaps have been left for two observers who did not participate (EGF and SH) to facilitate comparison with Figure 7.7.

strabismus (up to 29°) in such studies is too severe for them to be an appropriate model for many human strabismics, such as those studied here. In any case, the finding that binocular summation is intact in these subjects greatly reduces the set of possible model architectures considered in the following section.

There is an alternative arrangement which can produce summation ratios $\geq \sqrt{2}$. Tyler and Chen (2000) discuss a modern interpretation of probability summation, using a MAX operator within the framework of signal detection theory. They describe a specific case in which probability summation can produce substantial summation, around $\sqrt{2}$. However, this arrangement requires a linear contrast transducer, and therefore predicts that the slope of the psychometric function is shallow (Weibull $\beta \sim 1.3$). This is not the case here; the mean Weibull β values for the six amblyopes in the normalization experiment were 3.69 (good eye), 3.11 (bad eye), and 3.70 (binocular). This allows the probability summation explanation to be rejected, leaving physiological binocular summation as the only plausible account (see also Baker *et al.*, 2007d).

7.5 Models of amblyopia

Despite several studies on interocular interactions (Harrad and Hess, 1992; Levi *et al.*, 1980; Hood and Morrison, 2002; Pardhan and Gilchrist, 1992; Pardhan and Whitaker, 2000) and contrast discrimination (Ciuffreda and Fisher, 1987; Kiper and Kiorpes, 1994; Hess *et al.*, 1983) in amblyopia, a formal model has not yet been proposed to account for the abnormal empirical results. In this section, the Two Stage model (see Chapter 3 and Meese *et al.*, 2006) is modified in an attempt to account for the results of the above experiments.

Since the data of individual amblyopes are noisy, the averaged data, from Figures 7.2I and 7.3I, are used for modelling. It is also instructive to model the results from the normal observer with an ND filter over the left eye (Figure 7.1C&F). These two data sets are surprisingly similar in many respects, and are replotted in adjacent panels in Figure 7.9 for ease of comparison.

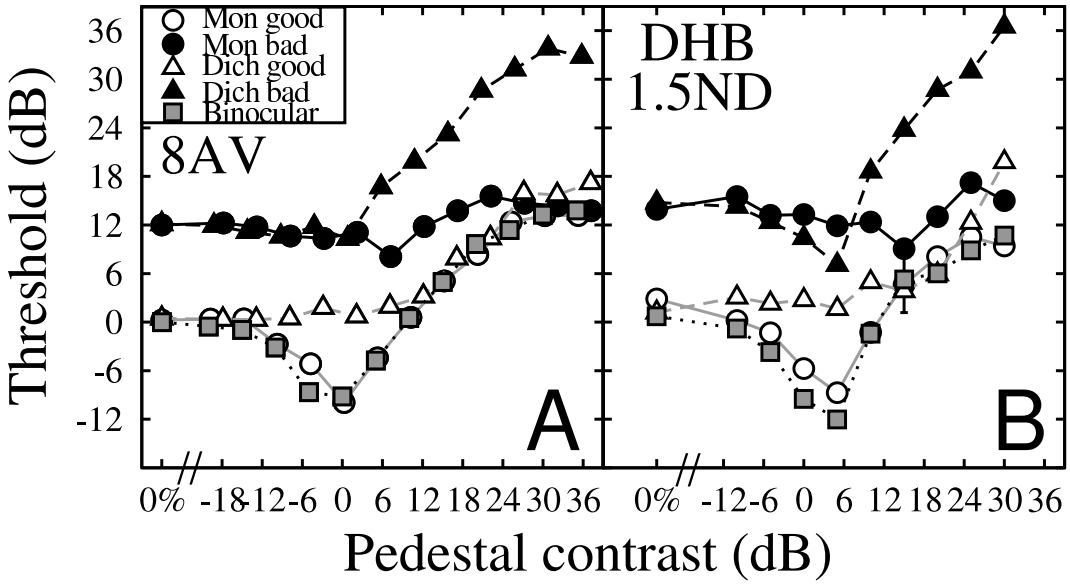

Figure 7.9: Amblyopic (A) and normal (B) contrast masking functions, plotted in adjacent panels to aid comparison. The data are replotted from Figures 7.1, 7.2 and 7.3. The neutral density filter produces the same pattern of results in a normal as are seen in amblyopes with no filter.

Several possible models are considered. Some of these models can be rejected *a priori*, on the basis of the experimental findings from the previous section. However, they are included for completeness, and because they represent conceptual models of the amblyopic visual system commonly discussed in the literature. It is important to develop such models computationally, so that they can be rejected, even if it is clear from the outset that they will not be viable. Strategies for modelling include removing binocular summation, increasing interocular suppression, inserting an attenuator at various locations, changing the model parameters, and the inclusion of stochastic (random) elements.

7.5.1 Independent channels

A potential model of amblyopia is one in which physiological binocular summation does not occur. In this model, the amblyopic system has two independent monocular channels which govern performance. The two stage model can be modified in this fashion by removing binocular summation between channels, such that the response of each channel is given by,

$$resp_L = \frac{\left(\frac{C_L^m}{S+C_L+\omega_D C_R}\right)^p}{Z + \left(\frac{C_L^m}{S+C_L+\omega_D C_R}\right)^q} \tag{7.1}$$

where ω_D is the weight of dichoptic masking, and all other terms are as used previously, with a similar expression for the right eye. The final model response is determined by the channel with the largest output (a peak-picker or MAX operator), which can also be implemented by Minkowski summation with a large exponent (see section 3.1.3). In principle, one monocular channel could then be attenuated (see next section), to produce the threshold elevation found for the bad eye.

This model arrangement predicts the absence of binocular summation at and above threshold which is observed empirically in amblyopia (squares in Figure 7.10 represent monocular and binocular performance). However, the dichoptic masking function has too steep a slope (grey line in Figure 7.10). This is because the test channel is not only suppressed by the mask (via divisive interocular suppression) but must also exceed it in activation in order to make good in the MAX operator. Less severe dichoptic masking can be achieved by setting $\omega_D = 0$. This removes the divisive interocular suppression and results in an entirely monocular model.

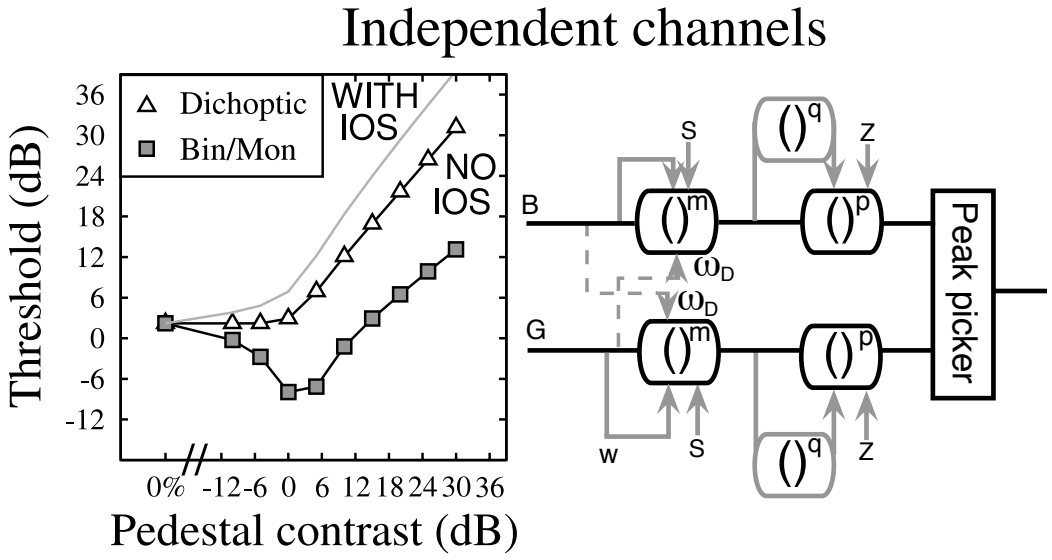

Figure 7.10: Behaviour and diagram of the independent channels model. Left and right channels are pooled using a peak picker at the output stage. Note that the binocular and monocular functions all superimpose onto the lower function (squares). The grey curve shows dichoptic masking with interocular suppression (ω_D=1 in equation 7.1), and the triangles are for dichoptic masking without interocular suppression (ω_D=0 in equation 7.1).

This model (shown diagrammatically in Figure 7.10) does produce levels of dichoptic masking consistent with those observed empirically (triangles in Figure 7.10). However, it

is hard to reconcile this arrangement with the finding that dichoptic masking from gratings of substantially different orientations appears normal in amblyopes (Harrad and Hess, 1992; Levi *et al.*, 1979, 1980). Contemporary studies of such effects (Baker *et al.*, 2007c; Baker and Meese, 2006a, and see Chapters 4 and 5) invoke some form of cross-channel interocular suppression, for which there is much physiological evidence (Sengpiel and Blakemore, 1996; Sengpiel and Vorobyov, 2005; Sengpiel *et al.*, 2006; Li *et al.*, 2005). If all binocular interactions are absent, including interocular suppression, it is not clear how dichoptic cross-channel masking could remain unaffected.

Two additional shortcomings of this model arrangement are that dichoptic facilitation is also lost, and there is no opportunity for empirical binocular summation to occur, even using a normalization procedure. Since dichoptic facilitation is apparent (Figure 7.5), and binocular summation has been shown for these observers after threshold normalization (see section 7.4), this configuration is clearly inadequate. However, it should be pointed out that the contrast discrimination data for two observers (EGF and JL), as well as those of Levi *et al.* (1979) are reasonably well described by this model.

7.5.2 Weight of interocular suppression

Another common explanation for sensitivity loss in amblyopia is that the amblyopic eye is suppressed by the normal eye by a greater amount than in normals. Thinking along these lines, Meese and Hess (2004) suggested that the differing slopes of dichoptic masking functions reported by Harrad and Hess (1992) could be modelled using different dichoptic weights in an early gain control equation similar to stage one of the Two Stage model. The results of the present experiment do not show such substantial deviations from a slope of one when the functions are displayed on normalized axes (see Figure 7.5). In some cases the masking function is shallower, usually in the bad eye, but it is not steeper for any of the subjects. This provides little opportunity to test this hypothesis using the present data set. However, it is interesting to explore the behaviour of the model when dichoptic weights $\neq 1$, given the pervasiveness of the suppression account of amblyopia.

The relationship between the magnitude of the suppressive weights and the strength of masking is a complex one. In the Two Stage model, dichoptic masking arises from two distinct processes. First, the signal in the test channel is directly reduced by the mask. This suppression acts in a straightforward, divisive manner, and accounts for the majority of the masking observed experimentally (see Chapter 6). However, the mask channel is also suppressed by the test signal. This means that for some combinations of mask and test, the overall activity of the model at the output stage (after binocular summation of mask and test) is less than with the same mask alone. This paradoxical property of the model was

explored more fully in Chapter 3.

The level of interocular suppression can be varied by introducing dichoptic weight terms onto the denominator of stage one, rewriting equation 3.6 as,

$$stage1 = \frac{{C_L}^m}{S + C_L + \omega_{DL} C_R} + \frac{{C_R}^m}{S + C_R + \omega_{DR} C_L}, \tag{7.2}$$

where ω_{DL} and ω_{DR} are the dichoptic weights in the left and right eyes respectively, and all other terms are as previously.

It is clear that altering the dichoptic weights will affect both of the suppressive processes discussed above, making the behaviour of the model difficult to predict intuitively. Figure 7.11 shows some example dichoptic functions for different values of the dichoptic weight terms. It is apparent that the interplay of the two processes results in dichoptic masking being approximately equal for each eye, even when the weights are very different. This means that varying the weights of the Two Stage model[3] in the manner suggested by Meese and Hess (2004) is not useful in describing either the present data set or the data of Harrad and Hess (1992), so long as binocular summation remains intact.

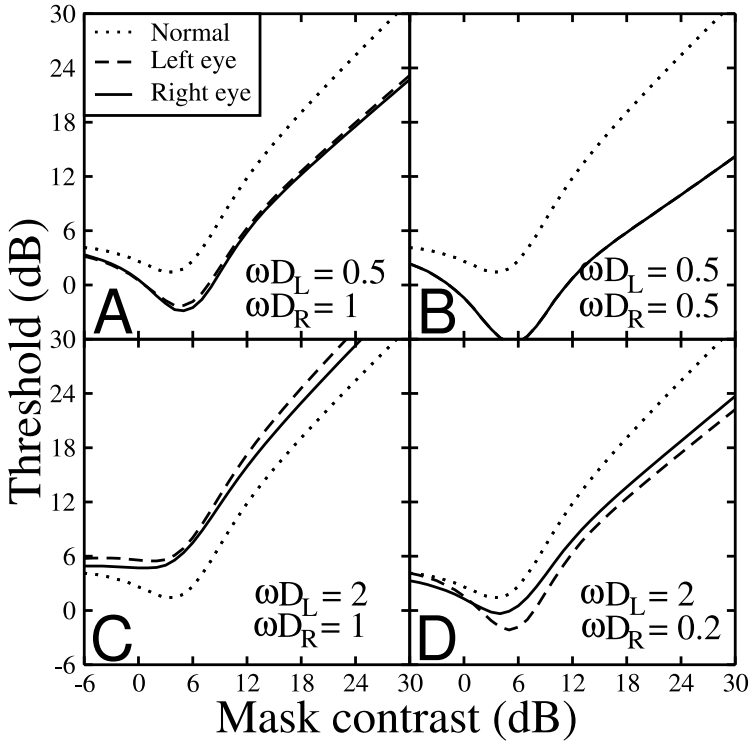

Figure 7.11: Dichoptic masking functions with different interocular weights. Dichoptic masking behaviour in left and right eyes is shown for a range of weight (ω_D) values. The dotted function gives normal behaviour for comparison, when both weights equal one.

[3]Note that this may not be the case for the model originally proposed by Meese and Hess (2004), however this model has already been rejected on other grounds (see Chapter 3 and Meese *et al.*, 2006)

7.5.3 Attenuator models

The simplest method of producing threshold elevation in the amblyopic eye is to insert an attenuator somewhere in the model. The attenuator divides any input by a fixed amount, the magnitude of which is an additional free parameter in the model. This must occur before binocular summation, in order that it does not also affect the good eye. In the Two Stage model, the obvious locations are before or after stage one (see diagrams in Figures 7.12 and 7.13).

When the attenuator is placed before stage one, the signal in the bad eye's channel, is equal to the input contrast divided by a constant:

$$C_B = \frac{C_L}{1+A} \tag{7.3}$$

where C_L is input contrast to the left eye, and A is a free parameter which determines the level of attenuation. This signal forms the input to the standard form of the Two Stage model, detailed in Chapter 3 and Meese *et al.* (2006).

When the attenuator occurs after stage one, it acts on the output of that stage, but before binocular summation. Thus, equation 3.6 becomes:

$$stage1 = \frac{\left(\frac{{C_L}^m}{S+C_L+C_R}\right)}{1+A} + \frac{\left(\frac{{C_R}^m}{S+C_R+C_L}\right)}{1}, \tag{7.4}$$

where symbols are all as previously (here, the left eye is assumed to be amblyopic).

Only three free parameters were used; k and S, which determine detection threshold and dip location, and A, which governs the amount of attenuation. All other parameters were fixed at the values from Meese *et al.* (2006). The best fits of these two models to the averaged data, and those of DHB for the ND filter condition, are shown in Figures 7.12 and 7.13. Parameters and error statistics are given in Table 7.3.

In general, the fits with the attenuator before stage one are reasonable (Figure 7.12). Thresholds are elevated in the bad eye, and the binocular function (squares and dotted curves) is correctly raised to around the level of the good eye's monocular function (upper panels, open circles and grey solid curves). However, the model fits poorly in two respects. First, the empirical monocular contrast discrimination function for the bad eye (filled circles) shows markedly weaker facilitation than either the model (black solid curves), or the equivalent function in the good eye (open circles; see also Figure 7.4). At higher pedestal levels the fit is also poor. Second, the model dichoptic masking function for the bad eye (lower panels, dashed black curves) is steeper than in the data (black triangles), indicating

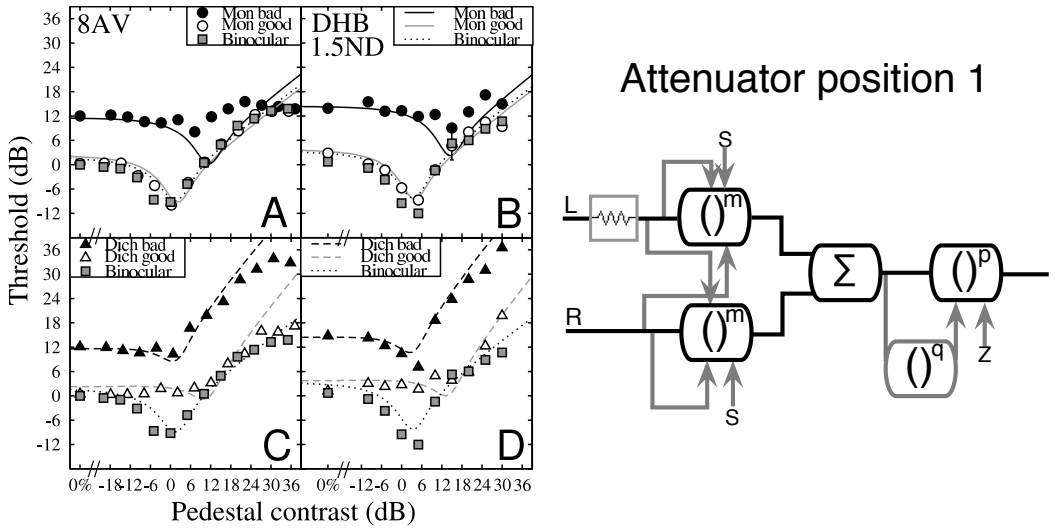

Figure 7.12: Best fits of the early attenuator model. Upper panels show monocular and binocular data, and lower panels show dichoptic and binocular data. The zig-zag icon in the model diagram indicates the attenuator, which occurs before stage one. Parameters and error statistics are given in Table 7.3.

Data set	Location	Figure	k	S	A	RMS error (dB)	AIC
Averaged	Before S1	7.12	0.22	0.87	1.92	3.98	95.78
DHB-ND	Before S1	7.12	0.21	1.20	2.40	2.72	56.03
Averaged	After S1	7.13	0.21	1.10	0.48	4.80	108.0
DHB-ND	After S1	7.13	0.19	1.51	0.66	3.95	74.69

Table 7.3: Parameters and error statistics for the attenuator models. Values are the best fits of the models, with three free parameters, estimated from 100 runs of a downhill simplex algorithm. The fits and model diagrams are shown in Figures 7.12 and 7.13.

that the model predicts stronger dichoptic masking than found empirically.

When the attenuator is placed after stage one, this results in an even poorer fit (Figure 7.13). Here, the dichoptic function becomes even steeper, and the model attempts to compensate by reducing the level of attenuation (the parameter A has lower values for this model, see Table 7.3), which results in a poorer fit at detection threshold. Clearly, this model is inferior to the early attenuator model and can therefore be rejected.

7.5.4 Varying the saturation constant

An early attenuator describes several features of the data well. However, it fails to capture the reduction in monocular facilitation seen in the amblyopic eye. One alternative change to the model which could address this is to introduce different values of the saturation constant, S, in the two eyes. The parameter S determines the point at which the transducer function passes from its accelerating region into its compressive region. Thus, it affects the horizontal placement of the dip regions, and also the baseline detection threshold. This makes different

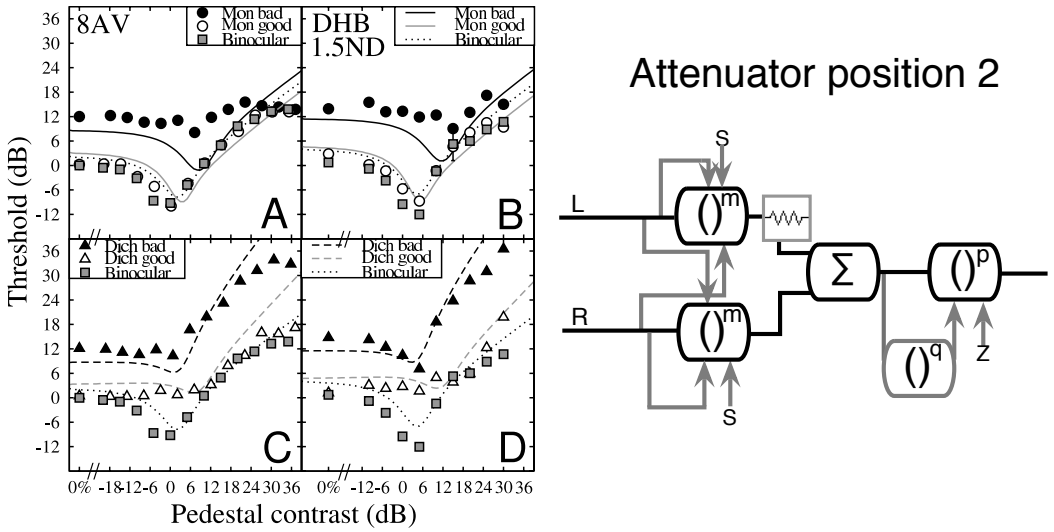

Figure 7.13: Best fits of the late attenuator model. Upper panels show monocular and binocular data, and lower panels show dichoptic and binocular data. The zig-zag icon in the model diagram indicates the attenuator, which occurs after stage one. Parameters and error statistics are given in Table 7.3.

Data set	Figure	k	S_G	S_B	A	RMS error (dB)	AIC
Averaged	7.14	0.24	0.61	6.09	-	4.16	98.66
DHB-ND	7.14	0.23	0.85	8.15	-	3.50	68.68
Averaged	7.15	0.23	0.70	2.24	0.88	3.77	94.31
DHB-ND	7.15	0.21	1.12	1.80	1.77	2.67	57.07

Table 7.4: Parameters and error statistics when S is estimated independently for the two eyes. S_G is the value for the good eye, and S_B is for the bad (amblyopic or filtered) eye. In the lower two rows, an attenuator is also included before stage one (see Figure 7.12). The fits are shown in Figures 7.14 and 7.15.

values of S across the eyes a plausible alternative to an attenuator.

Figure 7.14 shows the effects of allowing independent values of S in the two eyes (termed S_G for the good eye, and S_B for the bad eye). This results in RMS errors slightly larger than those for the best attenuator model from the previous section (see Table 7.4). Although producing a poorer overall fit, allowing S to vary does produce a good fit to the dichoptic masking function in the bad eye, as the function is somewhat shallower than in Figure 7.12. However, the binocular function is now elevated, above the level of the best eye's monocular function. This pattern is not consistent with the data of any of the subjects, although it might explain the binocular 'inhibition' at threshold (binocular summation ratios < 1) reported in the literature, and generally attributed to interocular suppression (Hood and Morrison, 2002; Pardhan and Gilchrist, 1992).

In Figure 7.15, the result of combining the early attenuator model with independent values of S is shown. This model has one extra free parameter (four in total), and provides marginally better fits than any of the previous models (see bottom rows of Table 7.4).

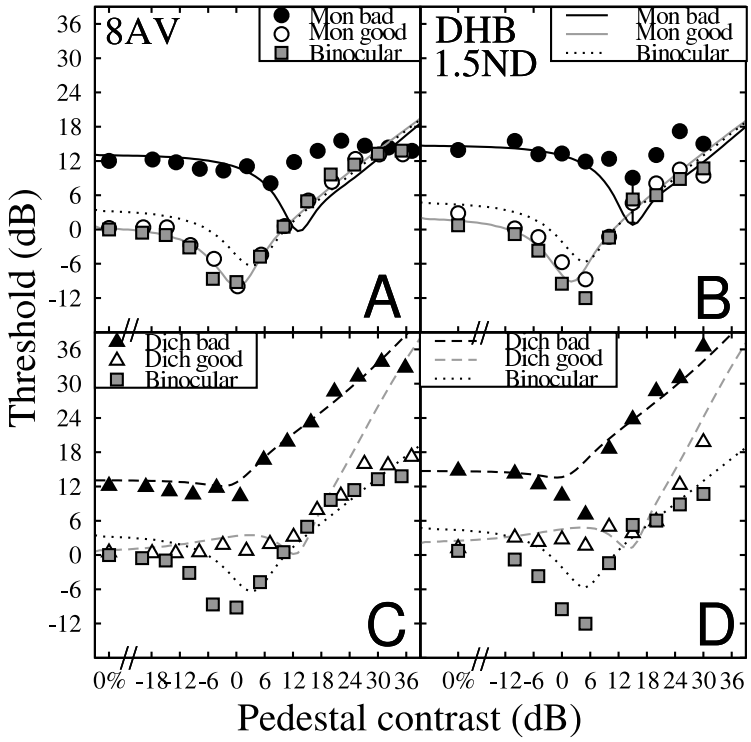

Figure 7.14: Model fits when values of S are independent in the two eyes. See Figure 7.12 for a description of panel layout. Parameters and error statistics are given in Table 7.4.

However, compared with the early attenuator model, the change in AIC score is negligible (-1.47 for Averaged data, +1.04 for DHB), suggesting that the inclusion of an additional free parameter is not warranted. Finally, the main shortcoming is still apparent - the monocular function in the bad eye is described very poorly by the model. The redundancy of this model is further supported by the following section, where an alternative model is presented which captures the pattern of monocular facilitation with the same number of free parameters.

7.5.5 Stochastic models

As demonstrated in Figure 7.4, monocular facilitaton was greatly reduced in the amblyopic eye for the majority of observers. This could come about if the transducer for the amblyopic eye was non-stationary (i.e. noisy). If the region of facilitation (or steepest gradient of the contrast response function) shifted slightly from trial to trial this would blur the dip region, making it broader and shallower, as found empirically.

Injecting additive noise into one monocular channel of the model can simulate this effect. The precise placement of the noise is important, as some exploratory modelling showed that noise occurring after stage one removes dips from all five functions. Including noise before stage one affects the monocular dipper function in the noisy eye, reducing the level of facilitation. It also slightly reduces the level of dichoptic facilitation in the good eye,

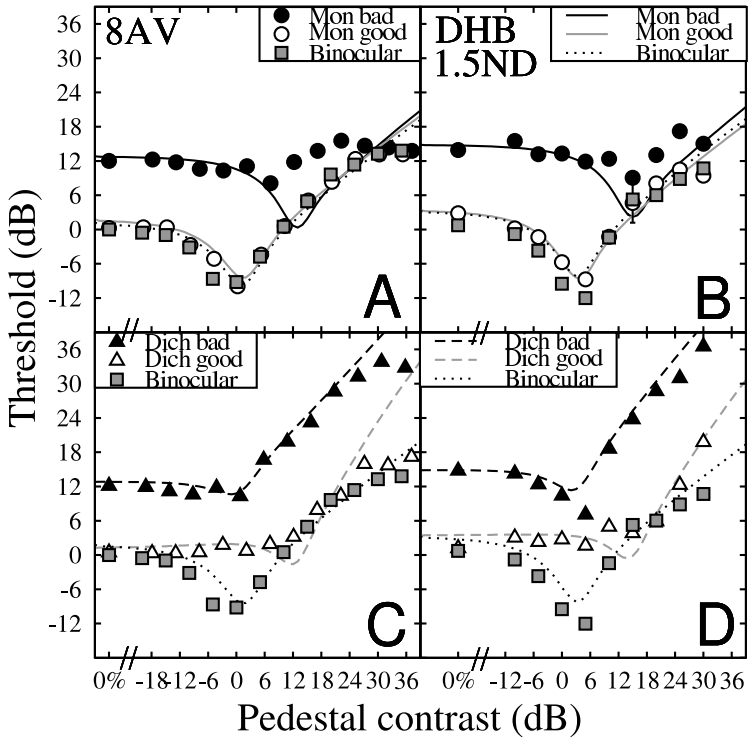

Figure 7.15: Model fits when values of S are independent in the two eyes and an attenuator is placed before stage one. See Figure 7.12 for a description of panel layout. Parameters and error statistics are given in Table 7.4.

consistent with Figure 7.5.

For the stochastic model, equation 7.3 becomes,

$$C_B = \frac{C_L + G_\sigma}{1 + A} \tag{7.5}$$

where G is a sample of zero mean gaussian noise, with standard deviation σ, and the attenuator (A) and all other terms are as previously. Because of the introduction of a stochastic term, a deterministic approach can no longer be used to calculate model predictions. Thus, the model was run on a trial-by-trial basis through the staircase procedure used for the experiments. Late additive noise, with a standard deviation proportional to k (see equation 3.13), was included on each interval of every trial, and the model decision was determined by the interval with the largest response. This procedure was repeated 2000 times, and an average threshold taken for each condition.

Model behaviour is shown in Figure 7.16 for σ=0.7, with all other parameter values taken from the first row of Table 7.3 and Meese *et al.* (2006). The stochastic model is very similar to the attenuator model, the main difference being the reduction of facilitation in the amblyopic eye. Only one additional parameter has been included (σ), resulting in a better

description of the data than the 3- and 4-parameter models from the previous sections. Furthermore, this model also predicts a further feature of the data: the increased noisiness, or variability, of thresholds in the amblyopic eye, as is now shown.

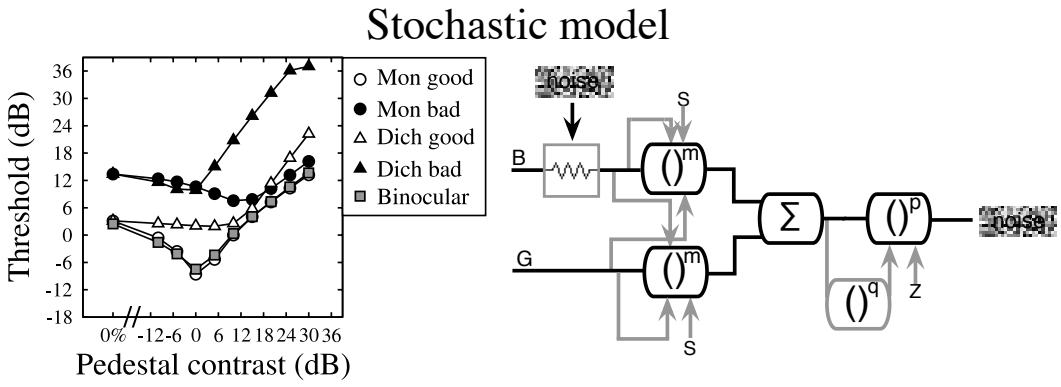

Figure 7.16: Behaviour and diagram of the stochastic model. The model behaves similarly to the attenuator model, the main difference being that monocular facilitation is weaker in the bad eye. This is caused by the early noise term, included before stage one (σ=0.7).

It is possible to quantify how 'noisy' the thresholds in the amblyopic eye are by comparing the standard error of the probit fit between good eye and bad eye. A larger standard error indicates noisier data. For the monocular data, a two-tailed paired samples t-test found the standard errors for the good eye to be significantly smaller than those for the bad eye ($t_{79} = -3.477, p < 0.001$) after a log-transform to remove positive skew. These data were further summarised by calculating the mean error ratio across subjects and contrast levels, which provides a convenient index to which model behaviour can be compared. The ratio of standard errors, $\frac{SE_{BAD}}{SE_{GOOD}}$, was calculated for each amblyopic observer at each pedestal contrast in the monocular condition. The geometric mean of these ratios (n=80) was 1.49, indicating that the probit standard error in the amblyopic eye tended to be greater than that in the normal eye.

This contrasts with what might be expected from what is known about the slope of the psychometric function in normal observers. The standard error of the probit fit is proportional to the standard deviation of the fitted cumulative gaussian. Thus, steeper psychometric functions tend to produce smaller standard errors. Psychometric slopes are typically steep in the region around detection threshold, and become shallower in and above the dip region due to 'small-signal linearization' by the pedestal (Bird *et al.*, 2002; Meese *et al.*, 2006, 2007a, and see Chapter 3). Since the monocular function for the amblyopic eye remains in the steep (pre-dip) region for a greater range of pedestal contrasts (see Figure 7.2), one might expect the slopes to be steeper on average, and thus produce smaller errors than for the good eye - the opposite to what was found. Thus, greater variability in the data from the amblyopic eye is evidence of additional noise in that pathway.

The same error ratio can be calculated for the stochastic model. With no early noise (σ=0), but an attenuator present (A=1.92), the mean error ratio was 0.52, consistent with the above discussion. When σ=0.7, the mean error ratio was 1.52 (averaged across 2000 independent simulations). Thus, the model accurately predicts the difference in standard error between amblyopic and normal eyes.

7.6 Discussion

An experiment was carried out to explore pedestal masking in amblyopia, using three ocular configurations of mask and test. The main findings were: i) dichoptic masking resembled that found in normals, ii) dichoptic facilitation was also observed, iii) monocular facilitation was weaker in the amblyopic eye than in the good eye, and iv) binocular summation was apparent when contrasts were normalized to threshold for each eye. The pattern of results from the amblyopic observers can be simulated adequately by placing a neutral density filter in front of one eye of a normal observer. Similarly, inserting an attenuator into one monocular channel of a computational model produced a good description of the data, which was further improved by including a stochastic term to increase the amount of noise in the amblyopic eye.

7.6.1 Comparison with previous studies

Kiper and Kiorpes (1994) studied contrast discrimination in experimentally strabismic monkeys. Their results have some similarities with the monocular dipper functions presented here. There is evidence for each of their three subjects that monocular facilitation is reduced in the amblyopic eye, particularly at higher spatial frequencies. The human data of Hess *et al.* (1983) focussed mainly on higher pedestal contrasts, so a comparison of facilitation for their subjects is not possible.

The most striking difference in findings is between the dichoptic data presented here, and those of Harrad and Hess (1992). Their study found consistently stronger dichoptic masking in the amblyopic eye of their strabismic subjects than would be expected for normals (i.e. a log-log slope of > 1). For the data reported here, dichoptic masking was either normal (on threshold normalized axes) or weaker than normal (see Figure 7.5). One possible explanation for this discrepancy is simply that the subjects tested by Harrad and Hess (1992) had a different pattern of deficits from those studied here. However, it seems unlikely that in a group of eight amblyopes, none would show the same behaviour.

A second possibility is that the difference in methodologies between the two studies could

be responsible for the disparate results. Harrad and Hess (1992) used full-field vertical gratings, presented using a mirror stereoscope. Given the difficulty in calibrating such devices, even with normal subjects, it is possible that the dichoptic stimuli might have become misaligned, resulting in a phase discrepancy between the eyes. However, given the results of Experiment 11 (Chapter 6), this would only be expected to produce *weaker* dichoptic masking than normal. A similar result would be expected if the stimuli were sufficiently misaligned to allow subjects to base their performance on some portion of the grating to the extreme left or right, which was not spatially coincident with the mask.

One further difference between studies is that, whereas horizontal gratings were used here, Harrad and Hess (1992) used vertical stimuli. There is some evidence that amblyopic deficits at threshold are greater for vertical than for horizontal stimuli (Sireteanu and Singer, 1980). This might be because vertical gratings are more important in the calculation of depth from binocular disparity, which is often abnormal in amblyopia (i.e. Ciuffreda, Levi and Selenow, 1990; Holopigian *et al.*, 1986). This is a plausible, and empirically testable explanation for the discrepancy, although ensuring correct alignment of vertical stimuli remains an issue with amblyopes, even when using shutter goggles. Furthermore, as discussed in section 7.5.2, it is not straightforward to encompass results such as those of Harrad and Hess (1992) within the framework of the Two Stage model.

A further comparison is between the data of the control subjects with neutral density filters, and the findings of Gilchrist and McIver (1985) and Gilchrist and Pardhan (1987). Whereas Gilchrist and colleagues report binocular inhibition (summation ratios < 1) for most subjects, even with weaker filters than the 1.5 log units used here, there is no strong evidence for this effect in the three normal subjects reported in Figure 7.6. This could be explained by subject differences, as half of the subjects reported by Gilchrist and McIver (1985) also failed to show the inhibition effect. Another possibility is that the use of 1.5 log unit ND filters has exceeded the luminance difference necessary to observe the inhibition effect. Given that binocular summation ratios are approximately similar ($\sim \sqrt{2}$) when the untested monocular eye is either patched or views mean luminance (see Table 1.1), it follows that the effect could only occur over a region of interocular luminance ratios lying between 0 and 1. It is thus conceivable that Gilchrist's effect only occurs over a limited range of luminance differences, up to around a log unit. However, the pilot experiments for this study used weaker ND filters, and also failed to find the effect for subject DHB.

7.6.2 Importance of neural binocular summation

The experiments reported here are the first strong evidence that the mechanisms of physiological binocular summation can remain intact in amblyopia. The presence of dichoptic

facilitation, and the normal levels of binocular summation found after compensating for the sensitivity difference between the two eyes, cannot be explained by any model which assumes independent channels for each eye (although for some amblyopes, such as the alternating esotrope of Harrad and Hess (1992), this architecture remains a possibility).

Besides reducing the possible set of model architectures for amblyopia, the preservation of binocular summation in strabismus suggests that binocular mechanisms may not be as prone to disruption as previously thought. Since the apparent lack of binocular summation has often been linked to the absence of stereovision in most amblyopes (Holopigian, Blake and Greenwald, 1988), this raises the question of whether stereopsis is truly absent in amblyopia, or whether a similar normalization method might reveal it to be undamaged (Halpern and Blake, 1988). Anecdotally, one subject in the present group (EMD) shows some limited stereopsis under normal testing conditions, and is also the subject who appears least 'amblyopic', when angle of strabismus, acuity loss, and sensitivity difference at threshold are all considered.

This hypothesis would certainly be worth testing, although it should be noted that stereopsis is reasonably robust with respect to interocular contrast differences, at least at the high spatial frequencies most affected in amblyopia (Hess, 1996; Legge and Gu, 1989; Schor and Heckmann, 1989; Cormack, Stevenson and Landers, 1997; Halpern and Blake, 1988).

7.6.3 Treatment of amblyopia

The similarity of results between amblyopic subjects, and normal subjects wearing ND filters in front of one eye is striking. As mentioned previously, similar findings have also been reported in a binocular rivalry paradigm (de Belsunce and Sireteanu, 1991; Leonards and Sireteanu, 1993). The usefulness of this as an experimental model in other paradigms, such as depth, motion and form perception, remains to be explored.

It has been shown that by equating the sensitivities of the two eyes, amblyopes show normal levels of binocular summation. Also, by placing a filter in front of one eye of a normal subject, the pattern of deficits found for amblyopic contrast vision can be simulated. Thus, it is possible that by placing an appropriately selected ND filter over the *good* eye of an amblyope, the two eyes might become equivalent, and the subject could behave normally in experiments such as those carried out here (albeit with an elevated contrast detection threshold).

This suggests a possible method of treatment for amblyopia in children, following surgery to correct strabismus. It is not known precisely what leads to the loss of stereopsis, or the

various other deficits observed in amblyopia. However, it seems plausible that it is the imbalance of inputs in the two eyes during development which is responsible. As noted above, mature stereopsis is relatively robust to interocular contrast differences at higher spatial frequencies (Hess, 1996; Legge and Gu, 1989; Schor and Heckmann, 1989; Cormack *et al.*, 1997; Halpern and Blake, 1988), but the developing system may be more delicate. Standard treatment, which involves patching the good eye for several hours at a time, would not be of use, as it never allows both eyes to function together. When the good eye is patched, it does not contribute a signal, and when it is functioning its signal will be much greater than that from the bad eye. Perhaps by equating sensitivities using a filter, the two eyes could provide the equal stimulation required for normal visual development.

Similar filters (called Bangerter foils) have occasionally been used in the treatment of amblyopia, but usually at a later stage, after patching (Rutstein, 2005). These can be attached to normal spectacle lenses, and produce less cosmetic impact than a patch. It is not clear whether the change in mean luminance caused by the ND filter is critical, or whether optical blurring or diffusion would produce the same result. These hypotheses could be easily tested in clinical trials.

7.6.4 Origins of the noise

The most successful model, which described the main features of the amblyope data, incorporated a noise term in the amblyopic eye, before the early gain control (stage one). This is consistent with other recent studies, which also find unusually high levels of noise in amblyopes (Levi and Klein, 2003; Levi, Klein and Chen, 2007; Huang, Tao, Zhou and Lu, 2007). However, it is not clear what this noise represents, or why is it present in amblyopia.

Firstly, it is worth pointing out that visual noise is not unique to amblyopes. There is a substantial literature on theories surrounding visual noise and its effect on performance in detection and discrimination paradigms (e.g. Watson, Borthwick and Taylor, 1997; Pelli, 1981, 1985; Rose, 1948; Barlow, 1962a,b; Henning and Wichmann, 2007). Noise is also fundamental to signal detection theory (Green and Swets, 1966), as it produces the psychometric function. Finally, small changes in detection thresholds over time have been reported (Wertheimer, 1953; Hallett, 1969a,b; Home, 1978), which could be responsible for the variability often reported in psychophysical tasks (i.e. non-stationarity of the psychometric function). Noise is undoubtedly present in normal subjects, perhaps from several sources, one of which appears to be greatly increased in amblyopia.

Noise recorded from individual neurones is generally multiplicative (it increases with signal strength) (Tolhurst *et al.*, 1981b), whereas that used in the model is additive. The appropriate type of noise for understanding psychophysical contrast discrimination behaviour is

a subject of much debate, and is beyond the scope of the present discussion (see Kontsevich *et al.*, 2002; Georgeson and Meese, 2006; Klein, 2006, for recent developments). However, it should be noted that using multiplicative noise at the early stage of the stochastic model did not produce convincing results (monocular dipper handles did not converge).

The increased noise could simply be an intrinsic property of the detecting neurones in visual cortex, as suggested by Levi and Klein (2003). These authors also suggest deficiencies in the subject's perceptual template, or stimulus uncertainty as alternative accounts of the increased noise they find. However, such concerns are less likely to affect the suprathreshold magnitude judgements in contrast discrimination experiments, as the appearance of the stimulus is known to the observer. Furthermore, one might expect an account based on a template model to affect both eyes equally, as presumably this is a high level perceptual judgement.

An alternative account of the noise is that it represents fluctuations in a standing level of masking from the other eye. The most likely cause of this would be masking caused by mean luminance, also known as 'zero-frequency masking' (Yang, Qi and Makous, 1995; Yang and Stevenson, 1999). This explanation has particular appeal, given that the data for the normal subject with an ND filter were also consistent with increased 'noise'. If this should prove correct, then the treatment proposed above would be dependent on the reduced luminance provided by an ND filter.

7.6.5 Conclusions

The work in this chapter has established that binocular interactions, particularly binocular summation, are intact in strabismic amblyopia. The substantial data set on contrast discrimination was used to successfully model amblyopic performance by developing the Two Stage model. Consistent with recent findings from the literature (Levi and Klein, 2003; Levi *et al.*, 2007; Huang *et al.*, 2007), the most successful model included greater noise in the amblyopic eye. The finding that amblyopia can be simulated by placing an ND filter in front of one eye of a normal subject has suggested a novel treatment for amblyopia, which remains to be tested clinically.

Chapter 8

Summary and Discussion

8.1 Summary of findings

The aim of this thesis was to investigate binocular interactions in human contrast vision. A series of experiments have explored interocular suppression, binocular summation, within- and cross-channel monocular effects, and amblyopic contrast vision. Each set of experiments has contributed to the development of a comprehensive model of binocular interaction, which provides a plausible architecture for the processing of contrast information in the binocular visual system.

Building on previous research, Chapter 3 investigated pedestal masking for monocular, binocular and dichoptic presentation. The results replicated and extended previous findings (Legge, 1979, 1984a). All functions showed facilitation at low pedestal contrasts and masking for higher pedestal contrasts. Masking in the dichoptic condition was particularly strong, conforming to Weber's law (log-log slope of 1; see Legge, 1979). Dichoptic facilitation, although weaker than in the other conditions, was also apparent, confirming occasional previous reports (Blake and Levinson, 1977; Levi *et al.*, 1979). Finally, binocular summation at threshold exceeded $\sqrt{2}$ for all subjects, contributing to a growing body of evidence (discussed in section 1.4.1) that previous accounts of binocular summation (Campbell and Green, 1965; Legge, 1984b) require revision.

The contrast discrimination data were used to evaluate four different models, two of which (the Two Stage model and the Twin Summation model) provided a good account of the results. A novel discrimination experiment using two pedestals of differing magnitudes, and a contrast matching experiment, provided further tests of the two models, both of which performed well. Finally, a prediction of the models, that non-monotonic psychometric functions can occur at high dichoptic mask contrasts, was tested experimentally and found

to be correct.

For stimuli with identical spatial properties (i.e. within-channel stimuli), both models were equally able to describe all available data. In the hope of differentiating between the models, further experiments investigated cross-channel masking, using mask and test stimuli which differed greatly in their spatial properties (Chapter 4). Such cross-channel masking has received considerable attention previously, but no attempt had been made to directly compare different modes of ocular presentation (i.e. monocular vs. dichoptic). The experiments used several ocular combinations of mask and test, and measured both contrast masking functions and masking as a function of stimulus duration. Crucially, monocular and dichoptic presentation produced masking of different magnitudes, which varied independently with stimulus duration, mask type and observer. This indicated that monocular and dichoptic masking must be caused by distinct processes, and ruled out models of masking in which either process occurred solely after binocular summation. The Twin Summation model was not suited to such an arrangement and was rejected, leaving the Two Stage model as the most viable model of contrast masking.

In concert with the psychophysical findings from Chapter 4, several recent single-cell studies have also explored cross-channel suppression in cat, and were reviewed in Chapter 5. These studies motivated four additional psychophysical experiments, which aimed to explore further the processes of cross-channel masking in humans. The first experiment measured dichoptic contrast discrimination in the presence of an additional (also dichoptic) cross-channel mask. This mask shifted the dipper function upwards and to the right, and was consistent with a model in which monocular and dichoptic suppression occurred in sequence. In the second experiment, masking was measured after first adapting to the mask. This reduced the level of dichoptic masking, implying a cortical locus for dichoptic suppression, consistent with single-cell findings in cat (Li *et al.*, 2005; Sengpiel and Vorobyov, 2005). Monocular masking was not reduced after adaptation, consistent with a solely pre-cortical locus (Freeman *et al.*, 2002; Li *et al.*, 2006; Priebe and Ferster, 2006).

As physiological studies have suggested that monocular and dichoptic suppression have different spatiotemporal profiles (Li *et al.*, 2005; Sengpiel and Vorobyov, 2005), the final two cross-channel experiments used orthogonal masks and tests which varied in spatial and temporal frequency. The first measured the tuning of masking for a 1cpd, 4Hz test grating, and masks of different spatial and temporal frequencies. Monocular masking was found to be broadly tuned along both dimensions, consistent with findings from physiology. Dichoptic masking was also broadly tuned, but showed a high spatial frequency cut which was not observed monocularly. The second experiment extended (using monocular and dichoptic presentation) a recent psychophysical study, which found that (binocular) orthogonal masking was strongly scale dependent (Meese and Holmes, 2007). Monocular masking followed

the same pattern as found binocularly, with masking being proportional to $\sqrt{TF/SF}$. Dichoptic masking, on the other hand, was spatiotemporally scale invariant. These findings support the distinct anatomical loci for monocular and dichoptic suppression reported by physiologists, as well as providing new insights into the processes of neural suppression in human.

In Chapter 6, a specific feature of the Two Stage model was tested, namely that dichoptic pedestal masking is caused by two processes, one of which depends on binocular summation. Using dichoptic masks presented in antiphase, it was possible to circumvent binocular summation, revealing the magnitude of masking caused by the other process. This was consistent with a zero-free-parameter prediction of the model, modified to incorporate separate phase channels, in which interocular suppression was phase insensitive, but binocular summation occurred only within a phase channel.

The antiphase paradigm was then used to explore the spatial frequency and orientation tuning of both interocular suppression and binocular summation. Operationally, binocular summation was found to be narrowly tuned for both orientation and spatial frequency, whereas interocular suppression had much broader tuning. This explained the unusual shape of dichoptic spatial frequency tuning functions reported by Legge (1979), which were shown to be the envelope of these two processes. The model developed to describe these data correctly predicted the shape of the dichoptic (orthogonal) spatial frequency tuning function from Chapter 5, indicating that it provides a good summary of a large stimulus space.

The models developed throughout most of the thesis were based on data from normal human observers. However, it is often instructive to consider cases in which a system is damaged or abnormal in some way. For binocular vision, amblyopia offers an interesting case, as amblyopes have no stereopsis, apparently lack binocular summation, and may suffer from increased interocular suppression of their amblyopic eye. In Chapter 7, the pedestal masking experiment of Chapter 3 was extended to strabismic amblyopes. Amblyopes were found to have normal dichoptic masking, and also some dichoptic facilitation. In the bad eye, the monocular dipper function shifted upwards and to the right, and showed weaker facilitation than in the normal eye. Binocular summation appeared to be absent at threshold using standard methods. However, such procedures do not take into account the sensitivity difference between the eyes. Using a summation paradigm, in which the contrast was threshold-normalized for each eye, normal levels of binocular summation were found for all amblyopes.

These findings rule out models based on two commonly held assumptions about amblyopia: that binocular summation does not occur (i.e. two independent monocular channels),

and that interocular suppression is increased. Instead, the results were well described by a version of the model which features an early attenuator, and increased noise in the amblyopic eye. Furthermore, the amblyopic deficits could be simulated by placing a neutral density filter in front of one eye of a normal observer. This provides a novel account of the contrast vision of strabismic amblyopes, as well as being the first attempt to model amblyopic contrast vision in detail.

The following section summarises the development of the two stage model throughout the thesis, and discusses some of the assumptions and limitations of the model, in the context of previous work. The remainder of this chapter highlights several trends in the preceding work, and discusses the findings in a broader context. Suggestions are made for further experiments and future research directions.

8.2 Model development

As outlined in the previous section, the two stage model has been substantially extended throughout this thesis in order to describe data from a variety of experiments. The 'full' version of the model offers a general description of all ocular configurations of pedestal masking (for both normal observers and strabismic amblyopes), as well as cross-channel masking and facilitation for all ocular configurations, dichoptic masking over a range of mask orientations and spatial frequencies, and adaptation to a monocular or dichoptic cross-channel mask. Figure 8.1 shows a schematic of this final version of the model, showing the loci of different masking effects

The full model shown in Figure 8.1 features two gain control stages, one before and one after binocular summation. At each gain control stage, the signal is subject to exponentiation, and divisive self-suppression. At the first gain control stage, within-channel dichoptic suppression also occurs, and is phase-insensitive. Divisive monocular cross-channel suppression, which is broadly tuned, takes place prior to stage one. Dichoptic cross-channel suppression occurs after stage one, and consequently, the masking signal has passed through stage one within its own detecting mechanism. As discussed in section 5.4.2, the cross-channel dichoptic signal may also impact at stage one, via a feedback process. Cross-channel facilitation occurs late, at stage two, and appears to be independent of spatiotemporal frequency (Meese and Holmes, 2007).

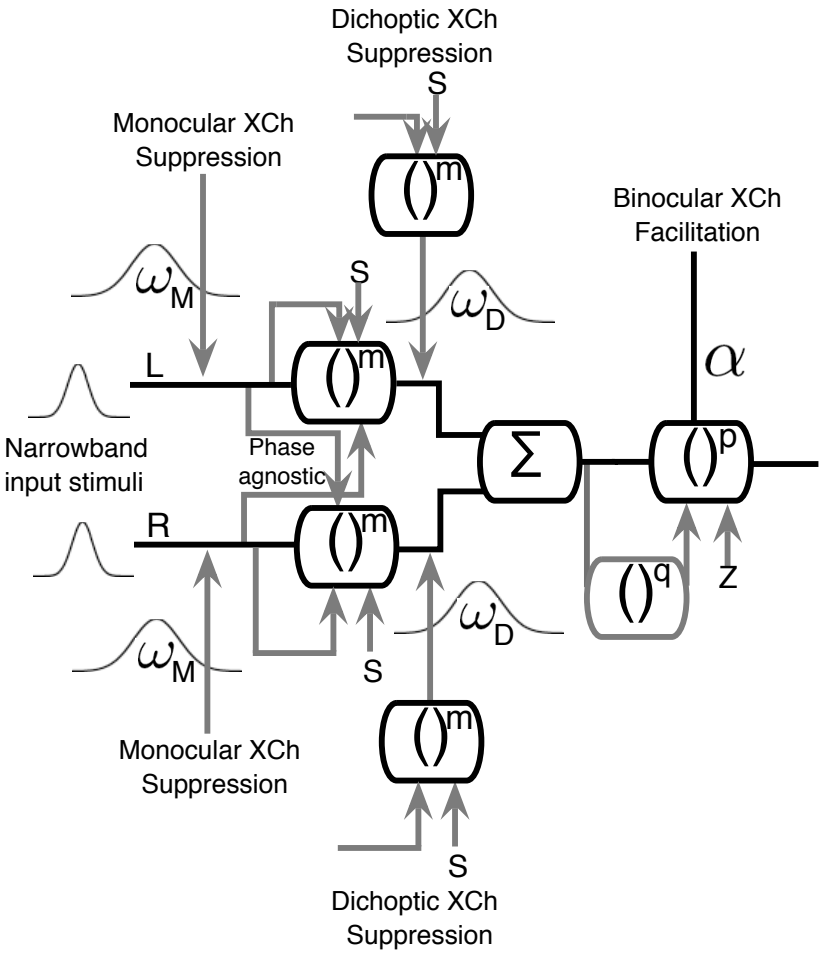

Figure 8.1: Diagram of full two stage model. Conventions are as for previous diagrams. The gaussian curves indicate orientation and spatiotemporal frequency tuning, for either the masking weights, or the excitatory input to the model. The cross-channel dichoptic suppressive pathway has passed through the first gain control stage within its own detecting mechanism, as indicated by the gain control box.

8.2.1 Interpretation of parameter values

The full version of the model contains a large number of parameters (at least ten). Comparing these to parameters from other models in the literature (i.e. Legge and Foley, 1980; Foley, 1994; Watson and Solomon, 1997) is difficult, as the two stage binocular architecture cannot be easily reconciled with previous models which do not consider eye-of-origin. For example, Foley's (1994) model produces a dipper function at a single gain control stage, with exponents of around 2.4 and 2 on the numerator and denominator respectively. To produce a comparable binocular dipper function, the two stage model utilises a shallow exponent (∼1.3) at stage one, and two much larger exponents (∼6) at stage two. This greater complexity allows the two stage model to be more flexible in describing the range of results found for different ocular combinations of mask and test, but makes parameter comparison with other models problematic. It is worth pointing out, however, that results from previous studies (using binocular presentation) can be well fit by the two stage model (i.e. Foley, 1994; Meese and Holmes, 2007; Bird *et al.*, 2002).

Stage two of the model features two exponents which take on unusually large values (p, $q > 5$) when fit using a simplex algorithm (Nelder and Mead, 1965). Since most cortical neurones do not exhibit such steep acceleration, these values perhaps deserve some comment. The second gain control stage could, in principle, be located at any point between binocular summation (V1) and the higher brain area at which 2AFC decisions are made. This suggests the possibility that stage two may represent the combined output of many stages of gain control (see Baker *et al.*, 2007a). Other components may contribute also, such as multiplicative noise (Kontsevich *et al.*, 2002; McIlhagga and Peterson, 2006) or stimulus uncertainty (Pelli, 1985; Petrov *et al.*, 2006). For these reasons, it would most likely be fruitless to search for neurones whose outputs are some higher power of their inputs. Instead, the exponents at stage two are treated as a mathematical convenience.

Tuning bandwidths

In Chapter 6, estimates were derived for the spatial frequency and orientation tuning of interocular suppression and binocular summation. The bandwidth estimates for binocular summation were very narrow indeed (<0.5 octaves, ∼6°), presumably because of the problems inherent in estimating bandwidths from spatially extensive stimuli when the mechanisms involved are phase-dependent (see Stromeyer and Klein, 1975; Wilson and Bergen, 1979; Bergen *et al.*, 1979). Indeed, early studies which used subthreshold summation to estimate channel bandwidths produced comparable estimates using large stimuli (Sachs *et al.*, 1971; Kulikowski *et al.*, 1973). This suggests that binocular summation occurs only within a channel, although the true bandwidth is probably much broader than that measured here,

because of the phase sensitivity of combination.

For interocular suppression, very broad bandwidths were found (>2 octaves, ∼50°). Since interocular suppression is apparently not phase dependent (see Chapter 6), it is unlikely that these estimates are affected by the problems discussed above. Indeed, there can be no physical interference between dichoptic gratings, as the mode of presentation dictates that any interactions between them must be neural. The estimates of interocular suppressive bandwidth for orientation are comparable to those reported for cat complex cells in layer 5/6 of V1 by Sengpiel *et al.* (2006, Figure 2). Interestingly, complex cells in layer 2/3 show isotropic tuning, suggesting that there may indeed be two distinct suppressive processes, implied by the additive constant (w) required in equation 6.3 to describe the dichoptic results. Equivalent single cell data are not available for spatial frequency tuning.

It is not clear how closely the estimates of suppressive bandwidth relate to the excitatory bandwidth of single neurones or psychophysical channels. Historically, masking was assumed to occur exclusively within a channel (Campbell and Kulikowski, 1966; Phillips and Wilson, 1984), so that the bandwidth of a masking function revealed the bandwidth of the detecting mechanism. However, more recent treatments have estimated narrower bandwidths by taking account of cross-channel suppression (see Meese and Holmes, 2003), indicating that the situation may not be so straightforward. This is particularly true of dichoptic masking, as the relationship between the bandwidths of a detecting neurone and the population of neurones which inhibit it (Heeger, 1992) could in principle be entirely arbitrary. Given that the spatial frequency bandwidth estimates for interocular suppression are substantially broader than excitatory bandwidths from both the psychophysics (Blakemore and Campbell, 1969; Georgeson and Harris, 1984; Bergen *et al.*, 1979) and physiology (Campbell and Robson, 1968; Campbell *et al.*, 1969; Maffei and Fiorentini, 1973; Movshon *et al.*, 1978; Tolhurst and Thompson, 1981; De Valois *et al.*, 1982a) literature (typically <2 octaves), a 1:1 mapping seems unlikely.

8.2.2 Neurophysiological loci of model components

Given the recent suggestion from the neurophysiology literature that monocular cross-channel suppression originates in the LGN, whilst dichoptic masking is a purely cortical process (see section 5.1), it is tempting to relate different parts of the model to specific brain areas. However, the complexity of the neural processes involved makes this problematic. The very first site, where monocular cross-channel masking occurs, is presumably located in the LGN (see section 5.1), and the later stages, from binocular summation onwards, are likely to be cortical. Intermediate stages, however, could either occur in the LGN (Bonin *et al.*, 2004), at the first synapse into cortex (Freeman *et al.*, 2002), at the site(s) of

binocular combination within V1 (Anderson and Movshon, 1989; Anzai, Ohzawa and Freeman, 1999a,b), or, most likely, at a combination of locations, perhaps incorporating feedback connections. How the various model components are implemented physiologically will hopefully become apparent with further single-cell work, but there is presently no compelling evidence to allow anatomical localisation of model stages.

8.2.3 Assumptions of the model, and alternative approaches

There are some assumptions implicit in the class of models used in this thesis. A reductionist approach is taken to many variables, such as stimulus strength (represented as a single number - the peak contrast) and stimulus area, which is never made explicit. Apart from the model discussed in section 5.4.2, all models considered have a feedforward architecture. Furthermore, model responses are typically based on the output of a single detecting mechanism, assumed to be optimally sensitive to foveal sinusoidal gratings of a particular spatial frequency and orientation. These assumptions allow the complex process of vision to be simplified into a few equations, which greatly expedites model development. At some level, such models must be appropriate for the stimuli used, as they provide a good account of much empirical data. However, there are some obvious failings, both specific, such as the unwanted 'peaks' discussed in section 6.4.1, and general - the model cannot produce a meaningful response to real images.

An alternative approach to modelling vision is to produce an image processing model, which takes a two dimensional greyscale (or, indeed, colour) image as an input. An influential model was developed along these lines by Watson and Solomon (1997), which was successful in modelling the psychophysical data of Foley (1994) and Foley and Boynton (1994). Such models are more elaborate that those developed here, and feature many mechanisms sensitive to specific spatial frequencies, orientations and locations in an image. Due to their complexity, simulations are computationally intensive, which increases development time. Furthermore, it is necessary to make many decisions regarding model parameterization, such as how many mechanisms should be included, how these should be pooled, what their bandwidths are, and how they interact with one another. Often such decisions can be informed by findings from the psychophysics and neurophysiology literature, or left as free parameters in the model.

Watson and Solomon's (1997) model has no knowledge of eye-of-origin. In principle, it could be extended to include two eyes, along the lines of the two stage model. Such a model might yield better fits to the dichoptic tuning data of Chapter 6, as it would have knowledge of the two dimensional phase relationships between mask and test. Furthermore, including multiple channels would permit off-channel looking in the orientation, spatial

frequency, and spatial domains, which may reduce the extra peaks in the manner discussed in section 6.4.1. Ultimately, a binocular image processing model must also be able to decode binocular disparity information in order to provide depth information - a domain distinct from the masking paradigm considered here. Hopefully, the present corpus of data will help to constrain the early stages of such a model, should one be constructed in the future.

8.3 Common trends across experiments

8.3.1 Binocular summation across experiments

Binocular summation ratios were measured at threshold in three experiments, on different observers and at different spatial frequencies, using both shutter goggles and a stereoscope. Figure 8.2 summarises the ratios, and presents a global average, which also appears in Table 1.1. The amblyopic subjects from Chapter 7 are not included, so the figure contains only data from normal subjects.

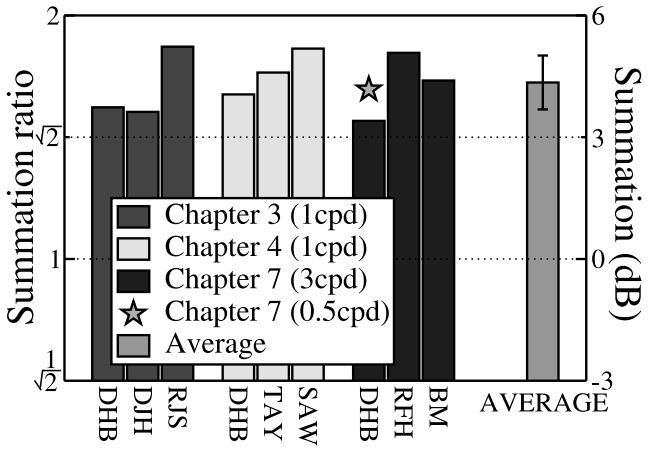

Figure 8.2: Binocular summation ratios from three experiments. Methodology for each experiment can be found in the appropriate chapter, given by the legend. The average (1.65) is the geometric mean, calculated across all the data points shown here (star included). Error bars are ± one standard deviation of the mean.

In contrast with early findings from the literature (reviewed in section 1.4.1), binocular summation ratios exceed $\sqrt{2}$ (1.41) for all observers, with the group average being 1.65. There is no obvious pattern across spatial frequency, however there do appear to be consistent individual differences. Subject DHB (author) shows relatively weak summation, with ratios around 1.5. Other subjects, such as RJS, SAW and RFH show much higher summation ratios, around 1.8. For subject RFH, this is consistent with results from two previous studies (Meese and Hess, 2004, 2005) in which high summation ratios were also measured for this subject. The cause of these individual differences in the level of binocular summation

is unknown.

These findings, as well as those of other recent studies (Meese *et al.*, 2006; Georgeson and Meese, 2005) have challenged a widely held view that binocular sensitivity is greater than monocular sensitivity by a factor of $\sqrt{2}$ (Legge, 1984a,b; Campbell and Green, 1965), as discussed previously in section 1.4.1. This finding has been incorporated into recent computational models (see Chapter 3 and Meese *et al.*, 2006) by including an early, almost linear transducer prior to binocular summation (see section 3.1.4). However, further modelling has shown that similar behaviour can be produced by aggregating the responses of large numbers of linear units, which differ in their degree of ocular dominance (Georgeson and Meese, 2007). Research into the processes of neural binocular summation is still ongoing.

8.3.2 Individual differences in dichoptic masking

Several experiments in this thesis have measured dichoptic masking using either within-channel (pedestal) masks, cross-channel (orthogonal) masks, or both. There is some indication from individual experiments (i.e. Experiment 13, Chapter 6) that the levels of orthogonal masking can vary greatly between different observers. Parallel (pedestal) masking, on the other hand, appears to be consistent across subjects (i.e. Experiment 1, Chapter 3). However, the results of most experiments have been plotted as absolute detection threshold against absolute mask contrast. This makes comparison between subjects difficult, as they rarely have comparable detection thresholds. A more appropriate manipulation is to normalize both axes to detection threshold, to allow the *relative* levels of threshold elevation to be compared.

Figure 8.3 shows maximum threshold elevations for nine subjects on normalized axes for both pedestal masks (panel A) and orthogonal masks (panel B). As expected, the pedestal data very closely follow the diagonal line of unity predicted both by Weber's law (Legge, 1979) and by the models of dichoptic masking described in Chapter 3. Thus, there are no clear individual differences in dichoptic pedestal masking. For orthogonal masking, however, there is a broader spread of points, varying over a 12dB range (a factor of 4). Clearly, any single masking function would describe these data poorly.

The individual differences in orthogonal dichoptic masking have no obvious cause. One possible explanation is that the level of masking is an indicator of some more general suppressive process in visual cortex. A recent study of area summation also found individual differences in contrast discrimination which were attributed to individual differences in suppression (Meese *et al.*, 2005b). Another possibility, that for high contrast dichoptic masks absolute thresholds are constant across observers, was ruled out in section 6.3.4.

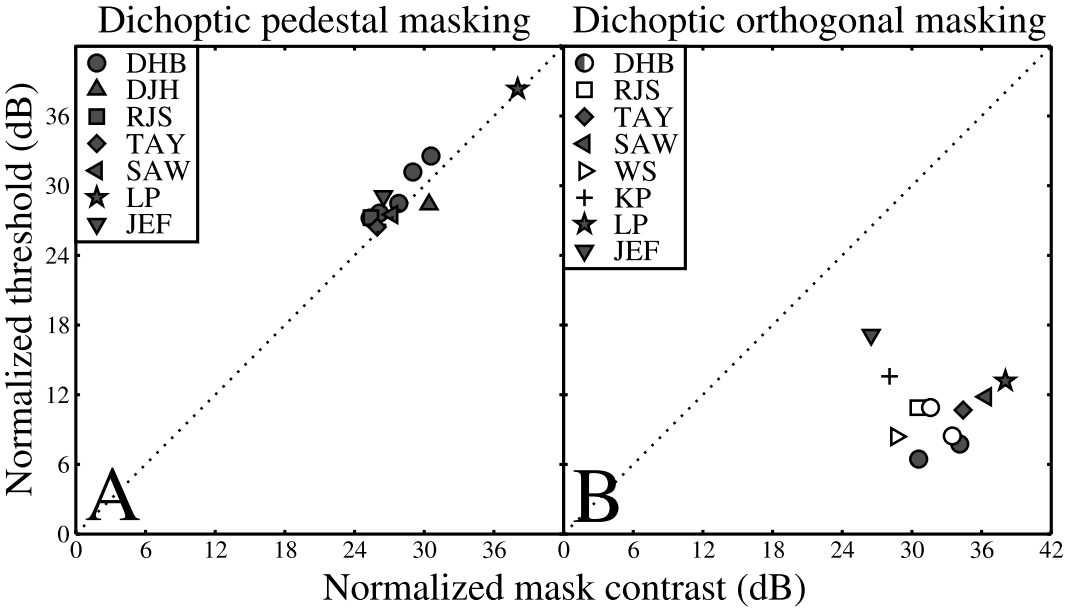

Figure 8.3: Dichoptic masking summarised across several experiments and observers, for pedestal masking (panel A) and orthogonal masking (panel B). Open symbols are for gratings presented with a temporal modulation, and filled symbols indicate static presentation (generally 200ms). If taken from a contrast masking (TvC) function, data are for the highest mask contrast used. Mask and test always had the same spatiotemporal properties, and varied only in contrast and orientation. Data are from Experiments 1, 5, 7, 9, 10, 13, and 14.

Individual differences in the levels of orthogonal dichoptic masking are reminiscent of other individual differences which have been reported using similar stimuli. In the binocular rivalry paradigm, orthogonal gratings are presented to opposite eyes for substantial periods, often several minutes at a time. This causes the observer's percept to alternate between the two gratings every few seconds. Individual differences in the average dominance duration (the length of time one stimulus is seen continuously) have frequently been observed across subjects (i.e. Pettigrew and Miller, 1998; Halpern, Patterson and Blake, 1987). Given the similarity between these paradigms, it seems plausible that dichoptic masking and binocular rivalry share a common neural circuitry, as has been suggested previously (i.e. Sengpiel *et al.*, 1995, 2001; Baker *et al.*, 2007c). However, further experiments would clearly be required to compare the two phenomena directly within a group of subjects, and thus reveal any common factor behind the individual differences.

8.4 Relation to other work

8.4.1 Classic psychophysics

As is to be expected for such a time-honoured research topic as binocular combination, many of the experiments summarised above draw on previous studies, and either confirm,

extend, or conflict with their findings. For example, in Chapter 7, a long-standing belief that strabismic amblyopes lack binocular summation was challenged, and found to be incorrect. Also, the dichoptic antiphase paradigm developed in Chapter 6 offers a new tool for exploring the tuning properties of binocular summation at suprathreshold levels. The meta-analysis of binocular summation ratios in Chapter 1 condenses and summarises much of the existing literature into a single table. The work here thus represents a step forward for many aspects of research into binocular contrast vision.

Much of the work in this thesis builds on the pioneering work of Gordon Legge. Legge explored dichoptic masking (Legge, 1979), binocular summation and integration (Legge and Rubin, 1981; Legge, 1984a,b) and, together with John Foley, introduced a class of models which have greatly influenced the contrast vision literature (Legge and Foley, 1980; Foley and Legge, 1981; Foley, 1994; Foley and Chen, 1997). This thesis offers a deeper insight into the mechanisms behind Legge's dichoptic masking data, particularly the spatial frequency tuning functions explored in Chapter 6. Convincing data on dichoptic facilitation, which historically has been somewhat elusive, have also now been gathered. Dichoptic facilitation is consistent with, and predicted by, the models of binocular contrast vision discussed in Chapter 3.

Importantly, a great deal more is now known about cross-channel masking, particularly regarding comparisons between monocular and dichoptic masking. With some notable exceptions (Legge, 1979; Meese and Hess, 2004), very few studies have directly compared these two types of masking psychophysically. This has hampered the development of binocular accounts of cross-channel masking, which must necessarily include both processes. The experiments here directly compared these two types of suppression, which has enabled development of model architectures.

One interesting realisation from these experiments is that binocular masking is not simply the sum of monocular and dichoptic effects. Instead, binocular masking more closely resembles monocular masking, in its tuning (Chapters 4 and 5), scale dependence (see section 5.6 and Meese and Holmes, 2007) and response to adaptation (see section 5.3 and Foley and Chen, 1997). Careful experimentation and modelling (see sections 4.2.3 and 4.5.2) suggests that this is because during binocular presentation the two masks (one in each eye) suppress each other dichoptically. This results in weaker masking of the test from the dichoptic cross-channel pathway, so this form of suppression contributes much less to threshold elevation. Monocular masking, which occurs early, is unaffected by this dichoptic suppression of masks, and therefore dominates during binocular presentation.

8.4.2 Neurophysiology

Recently, several single-cell studies have focussed on cross-channel effects, largely motivated by the suggestion (Freeman *et al.*, 2002) that cross-channel suppression may not be an intracortical phenomenon, as was previously believed (Morrone *et al.*, 1982). This work enables several important comparisons to be drawn between physiology and psychophysics. Reassuringly, there is close agreement between the two methodologies, which allows conclusions drawn from cat and primate physiology to be extended to humans with reasonable confidence. There is converging evidence from both approaches to suggest that monocular suppression occurs pre-cortically, possibly in the LGN, and dichoptic suppression occurs later in visual cortex (see Chapter 5).

In addition to this, some of the findings from this thesis suggest experiments which could be performed physiologically. The most striking example is the contrast between the scale-invariance of dichoptic masking and the strong scale-dependence for monocular and binocular cross-channel masking (Meese and Holmes, 2007). An understanding of the neural underpinnings of these phenomena is key to developing accurate computational models, and may provide insights into the mechanisms used by the visual system to encode apparent speed (see Rainville *et al.*, 2005; Hammett, Champion, Morland and Thompson, 2005; Harris, 1980). Such research might also provide insight into the mechanisms of clinical suppression, binocular rivalry, and diplopia.

8.4.3 Ding & Sperling

During the course of this work, another body of research has been published concerned with understanding binocular combination of contrast. Ding and Sperling (2006a,b) also developed a gain control model of binocular combination (see also Sperling and Ding, 2006). Their model explains results from a paradigm in which gratings differing slightly in phase and contrast are presented dichoptically. The cyclopean image formed by these stimuli is then probed using an edge location task in order to determine the contribution each stimulus makes to the binocular neural representation. Several variations on this paradigm are explored, including the addition of orthogonal masking gratings and bandpass filtered noise.

Ding and Sperling's model differs in a number of ways from the models developed here. Since their experiments do not assess the contrast response of monocular or binocular mechanisms, there is no provision in their model for the contrast detection and discrimination paradigms studied in Chapter 3. Furthermore, the Ding & Sperling model assumes a single cyclopean channel which sums across monocular phases algebraically. This assumption is

not compatible with findings from experiments using antiphase gratings (particularly Simmons, 2005), as the model predicts cancellation when contrasts are equal. The Two Stage model, on the other hand, assumes several independent binocular phase channels, so is able to account for antiphase summation data (see Chapter 6).

Despite these differences, there are some substantial similarities between the early stages of the Ding & Sperling model and the Two Stage model, as both incorporate divisive interocular suppression via a gain control process, prior to binocular summation (see Figure 8.4). It is encouraging that two rather similar model architectures can emerge from very different experimental paradigms.

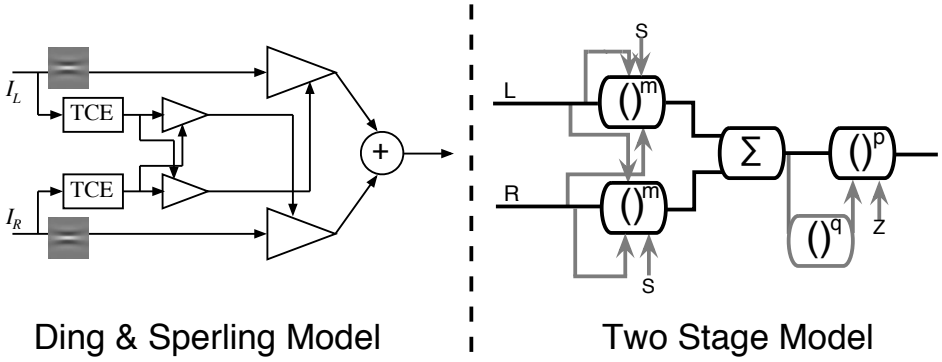

Figure 8.4: Comparison of the Ding & Sperling and Two Stage models. Pictured on the left is the model proposed by Ding and Sperling (2006a,b). The boxes labelled 'TCE' indicate total contrast energy, which is a weighted sum of all orientation and spatial frequency channels. The gabor patches signify oriented filtering of inputs (I_L and I_R) to the two eyes. The triangles are gain control stages, which divide their input (arrow at the left of the triangle) by the divisive signal at the diagonal edge, plus unity. The circle at the right of the diagram indicates algebraic binocular summation of inputs. The Two Stage model (right) is as described in section 3.1.4.

8.4.4 Rivalry models

Another close similarity can be drawn between the Two Stage model, and recent multistage models of binocular rivalry (i.e. Wilson, 2003; Freeman, 2005). Again, such models are not concerned with the pedestal effects and brief durations of masking explored here. However, there is a clear architectural similarity between such models and the Two Stage model. It seems likely that if dichoptic masking and binocular rivalry share a common mechanism (Sengpiel *et al.*, 1995), stage one of the Two Stage model could map onto the first stage of a multistage rivalry model. Indeed, stage one is very similar to the interocular inhibitory stage of Lehky's (1988) model.

The main distinction between masking and rivalry models is that the former are typically implemented as feedforward architectures (though see section 5.4.2), whereas the latter use feedback processes. Schematic diagrams of three binocular rivalry models are shown in

Figure 8.5, and can be compared to the two masking model architectures shown in Figure 8.4. All five models feature early interocular suppression. Interestingly, as rivalry models have evolved, they have included additional stages of suppression between channels. In the Freeman (2005) model, the distinction between monocular and binocular stages is less important than in the other models, as there are potentially more than two competitive stages. This means that the model might also be applicable to other bistable or competitive neural processes (i.e. the Necker cube).

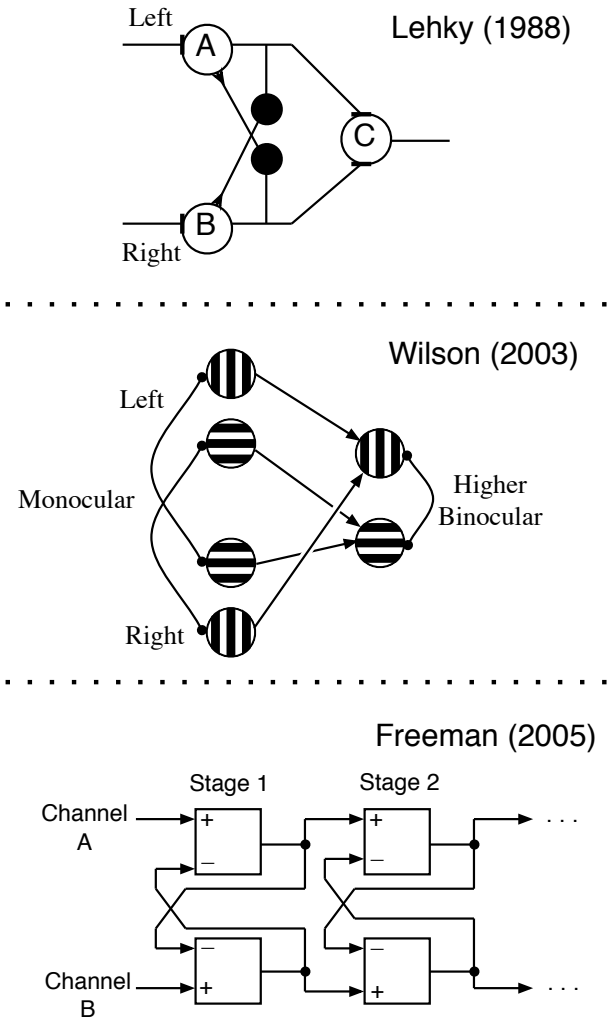

Figure 8.5: Schematic diagrams of three binocular rivalry models. All diagrams proceed from monocular inputs at the left, through to later (binocular) stages on the right. The Lehky (1988) model (top) features a single stage of interocular inhibition, before the monocular neurones (A, B) converge onto a binocular cell (C). The Wilson (2003) model (middle) features inhibition at both monocular and binocular stages, depicted by the ball-ended lines. Grating icons indicate the orientation of stimuli which each channel responds to. Finally, the Freeman (2005) model (bottom) comprises an indefinite number of identical stages, with inhibition between channels occurring at each.

A model which incorporated aspects of both rivalry and masking models would be extremely powerful, as it might help to reconcile performance measures (masking) with percep-

tual metrics (rivalry) over a wide range of presentation durations and stimulus configurations. Such a model might also allow for a quantitative comparison of the contributions of low level suppression, and higher level (contextual, or top-down) effects in rivalry. There is currently much controversy in the rivalry literature about the relative importance of these two processes (Blake and Logothetis, 2002; Alais and Blake, 2005). However, developing a model along these lines would require substantial further work, both empirical and computational.

8.5 Future directions

8.5.1 Binocular summation: bandwidth and luminance dependence

One conclusion drawn from the experiments in Chapter 6 is that the bandwidth of binocular summation is narrow when using large patches of grating ($\sim 15^\circ$ orientation, $< \frac{1}{2}$ octave spatial frequency). While similar conclusions can be drawn from other studies (Bacon, 1976; Kertesz and Jones, 1970; Blakemore, 1970), these typically also use large stimuli, and so are subject to an underestimation of bandwidth discussed in section 6.6 (see also Bergen *et al.*, 1979). More work is required to measure the bandwidth of summation accurately, using appropriate (spatially restricted) stimuli, and perhaps additional paradigms such as subthreshold summation.

A further unresolved issue is the importance for binocular summation of the mean luminance level presented to the untested eye, when recording a monocular threshold. As noted in Chapter 3, it appears that patching or occluding the untested eye results in slightly lower summation ratios than does presenting it with mean luminance. Whether this is a reliable effect when using the same group of subjects and a modern psychophysical procedure (many early studies which used patches also used the method of adjustment) remains to be seen.

8.5.2 Masking effects in other stimulus domains

The experiments in this thesis have focussed entirely on achromatic, luminance-defined grating stimuli. However, recently there has been some exploration of binocular contrast vision using chromatic stimuli (Simmons and Kingdom, 1998; Simmons, 2005; Medina and Mullen, 2007). Thus far, few experiments have been done, and the results have not lead to a specific model proposal. However, one finding of particular interest is that orthogonal masking using chromatic (isoluminant) stimuli is both strong and scale-invariant (Medina *et al.*, 2007). Isoluminant stimuli excite only neurones in the parvocellular stream (p-cells); the same cells which appear to produce weak masking for luminance-defined (achromatic) stimuli (see section 5.6, and Meese and Holmes, 2007). These findings therefore present a

puzzle for the m- and p-cell account of speed effects in cross-channel masking discussed in section 5.6.4. One likely explanation is that the scale-invariance found for both dichoptic and isoluminant stimuli points to a common neural locus, which other studies of dichoptic suppression (reviewed in section 5.1) indicate is cortical.

This account suggests that cortical suppression exists to provide suppression for stimuli which are not subject to the magnocellular, pre-cortical (and monocular) process. Such stimuli should include dichoptic and isoluminant stimuli, but also annular masks and 'slow' (meaning high spatial, low temporal frequency) stimuli. Indeed, there is evidence that the masking produced by annuli has a very different character to that elicited by superimposed masks, showing a different time-course physiologically (Webb *et al.*, 2005), and evidence of saturation psychophysically (Petrov *et al.*, 2005; Challinor, Meese and Summers, 2007). Furthermore, a recent neurophysiological study has found evidence for cortical suppression of low temporal frequency monocular stimuli (Sengpiel *et al.*, 2006), which are presumably 'missed' by the earlier pre-cortical process. However, there is need for a series of targeted experiments to directly compare all of these potential causes of cortical suppression.

Whereas the effect of using different temporal envelopes on cross-channel dichoptic masking has been fruitful, there has been no equivalent investigation for within-channel stimuli. There is evidence that both the dip and handle regions of binocular dipper functions (Wichmann, 1999; Phillips and Wilson, 1984; Georgeson and Georgeson, 1987; Lehky, 1985; Legge, 1978; Maehara and Goryo, 2005), and also binocular summation at threshold (Rose, 1980; Georgeson and Meese, 2005), are affected by temporal envelope and stimulus duration. However, dichoptic masking always conforms to Weber's law (log-log slope of 1), even when tested over diverse spatial frequencies (Legge, 1979) and temporal conditions (compare the data of Meese *et al.*, 2006, and Maehara and Goryo, 2005). There is scope within the Two Stage model to vary the slope of monocular and binocular dipper handles (both of which are defined by the exponents p and q) while keeping the dichoptic slope constant. However, it remains to been seen empirically whether the Weber's law behaviour holds over all spatiotemporal frequencies.

Finally, there has been much recent psychophysical work on the effects of masking stimuli which differ from the target in spatial position (i.e. Meese *et al.*, 2007b; Petrov *et al.*, 2005; Petrov and McKee, 2006; Huang, Hess and Dakin, 2006; Cass and Spehar, 2005). Some studies have begun to explore dichoptic effects (Petrov and McKee, 2006; Huang *et al.*, 2006; Huang, 2007, present author, unpublished observations), and have generally found that masking remains intact whereas facilitation is abolished. However this has not been investigated exhaustively, or in light of the results presented here.

8.5.3 Cross-channel masking in amblyopia

An important result from Chapter 5 is that monocular and dichoptic orthogonal masking have very different scale dependencies. Dichoptic masking is scale-invariant, being equal at all spatiotemporal frequencies, whereas the magnitude of monocular masking is determined by the square root of stimulus speed. This lawful but contrasting pattern of results has potential relevance to the study of amblyopia. Dichoptic suppression is believed to be cortical in origin, and the deficits observed in amblyopia are often attributed to damaged visual cortex (Sengpiel *et al.*, 2006). If this is the case, then amblyopes might be expected to show abnormalities in dichoptic masking. However, an alternative account of amblyopia is that there is damage to the LGN (Ikeda and Wright, 1976; von Noorden and Crawford, 1992; Chino, Cheng, Smith, Garraghty, Roe and Sur, 1994), in which case dichoptic (cortical) suppression may be largely unaffected. Some previous research has suggested that dichoptic cross-channel masking is normal in amblyopia (Harrad and Hess, 1992; Levi *et al.*, 1980). However, the conditions in those two studies were run on only one subject each, and covered too small a range of conditions to be conclusive. The lawful (scale-invariant) dichoptic masking, and strongly scale-dependent monocular masking demonstrated in section 5.6 offer a strong test of these competing hypotheses. If amblyopes show deviations from scale-invariance for dichoptic masking, this would be evidence in favour of a cortical deficit. However, if the scale dependency of monocular masking (see also Meese and Holmes, 2007) is disturbed, this would indicate a pre-cortical (i.e. LGN) deficit. It is also possible that amblyopes will show abnormal behaviour for both, or neither, types of masking.

8.6 Conclusions

The experiments in this thesis have contributed to a growing body of research on the binocular visual system. Several classic findings from the literature have been re-evaluated in light of new evidence, and much new ground has been covered. Of particular note, the early architecture of cross-channel masking is now better understood, the tuning of interocular suppression has been measured, and a plausible model of contrast vision in amblyopia has been developed.

As mentioned previously, much of the work has already been published in article (Baker *et al.*, 2007a,c,d; Baker and Meese, 2007; Meese *et al.*, 2006) and abstract (Baker *et al.*, 2005, 2007e,f; Baker and Meese, 2006a,b; Georgeson *et al.*, 2005; Meese *et al.*, 2005a) format, as well as being presented at national and international conferences. Furthermore, these publications form part of a corpus of research on suppression and summation in human vision, supported by an EPSRC grant, encompassing many areas of spatial vision (see Meese

and Holmes, 2007; Meese and Summers, 2007; Meese *et al.*, 2005b, 2007a,b). It is hoped that this collection of work will stimulate further research into human contrast vision.

Bibliography

Akaike, H. (1974). A new look at the statistical model identification, *IEEE Transaction on Automatic Control* **19**: 716–723.

Alais, D. and Blake, R. (2005). *Binocular Rivalry*, Bradford.

Albrecht, D. G. and Geisler, W. S. (1991). Motion selectivity and the contrast-response function of simple cells in the visual cortex, *Vis Neurosci* **7**: 531–546.

Albrecht, D. G. and Hamilton, D. B. (1982). Striate cortex of monkey and cat: contrast response function, *J Neurophysiol* **48**: 217–37.

Albrecht, D. G., Farrar, S. B. and Hamilton, D. B. (1984). Spatial contrast adaptation characteristics of neurones recorded in the cat's visual cortex, *J Physiol* **347**: 713–39.

Anderson, P. A. and Movshon, J. A. (1989). Binocular combination of contrast signals, *Vision Res* **29**: 1115–32.

Anzai, A., Ohzawa, I. and Freeman, R. D. (1999a). Neural mechanisms for processing binocular information i. simple cells, *J Neurophysiol* **82**: 891–908.

Anzai, A., Ohzawa, I. and Freeman, R. D. (1999b). Neural mechanisms for processing binocular information ii. complex cells, *J Neurophysiol* **82**: 909–24.

Arditi, A. R., Anderson, P. and Movshon, J. (1981). Monocular and binocular detection of moving sinusoidal gratings, *Vision Res* **21**: 329–336.

Asper, L., Crewther, D. and Crewther, S. G. (2000). Strabismic amblyopia part 1. psychophysics, *Clin Exp Optom* **83**: 49–58.

Bacon, J. H. (1976). The interaction of dichoptically presented spatial gratings, *Vision Res* **16**: 337–44.

Baker, D. H. and Meese, T. S. (2006a). Cross-orientation suppression occurs before binocular summation: evidence from masking and adaptation, *Journal of Vision* **6**: 821.

Baker, D. H. and Meese, T. S. (2006b). Monoptic and dichoptic cross-orientation masking are not the same mechanism, *Perception* **35**: 421.

Baker, D. H. and Meese, T. S. (2007). Binocular contrast interactions: dichoptic masking is not a single process, *Vision Res* **47**: 3096–3107.

Baker, D. H., Meese, T. S. and Georgeson, M. A. (2005). Contrast discrimination with simultaneous monocular and dichoptic masks, *Perception* **34**(S): 202.

Baker, D. H., Meese, T. S. and Georgeson, M. A. (2007a). Binocular interaction: contrast matching and contrast discrimination are predicted by the same model, *Spatial Vision* **20**: 397–413.

Baker, D. H., Meese, T. S. and Patryas, L. (2007b). Binocular summation is more tightly tuned to spatial frequency, orientation and spatial phase than interocular suppression, *Perception* **36**: 1401.

Baker, D. H., Meese, T. S. and Summers, R. J. (2007c). Psychophysical evidence for two routes to suppression before binocular summation of signals in human vision, *Neuroscience* **146**: 435–48.

Baker, D. H., Meese, T. S., Mansouri, B. and Hess, R. F. (2007d). Binocular summation of contrast remains intact in strabismic amblyopia, *Invest Ophthalmol Vis Sci* **48**: 5332–5338.

Baker, D. H., Meese, T. S., Mansouri, B. and Hess, R. F. (2007e). Monoptic, dichoptic and binocular masking in strabismic amblyopia, *Perception* **36**: 302.

Baker, D. H., Meese, T. S., Patel, K. and Sarwar, W. (2007f). Interocular suppression is scale invariant, but ipsiocular suppression is weighted by flicker speed, *Perception* **36(S)**: 60.

Barbeito, R., Levi, D. M., Klein, S., Loshin, D. and Ono, H. (1985). Stereo-deficients and stereoblinds cannot make utrocular discriminations, *Vision Res* **25**: 1345–8.

Barlow, H. B. (1962a). Measurements of the quantum efficiency of discrimination in human scotopic vision, *J Physiol* **160**: 169–88.

Barlow, H. B. (1962b). A method of determining the over-all quantum efficiency of visual discriminations, *J Physiol* **160**: 155–68.

Bergen, J. R., Wilson, H. R. and Cowan, J. D. (1979). Further evidence for four mechanisms mediating vision at threshold: sensitivities to complex gratings and aperiodic stimuli, *J Opt Soc Am* **69**: 1580–7.

Birch, E. E. (1979). *Binocular processing of contrast*, PhD thesis, US Santa Barbara.

Bird, C. M., Henning, G. B. and Wichmann, F. A. (2002). Contrast discrimination with sinusoidal gratings of different spatial frequency, *J Opt Soc Am A Opt Image Sci Vis* **19**: 1267–73.

Bjorklund, R. A. and Magnussen, S. (1981). A study of interocular transfer of spatial adaptation, *Perception* **10**: 511–8.

Blackwell, K. T. (1998). The effect of white and filtered noise on contrast detection thresholds, *Vision Res* **38**: 267–80.

Blake, R. and Cormack, R. (1979). On utrocular discrimination, *Perception & Psychophysics* **26**: 53–68.

Blake, R. and Fox, R. (1973). The psychophysical inquiry into binocular summation, *Percept Psychophys* **14**: 161–185.

Blake, R. and Levinson, E. (1977). Spatial properties of binocular neurones in the human visual system, *Exp Brain Res* **27**: 221–232.

Blake, R. and Logothetis, N. K. (2002). Visual competition, *Nat Rev Neurosci* **3**: 13–21.

Blake, R. and Rush, C. (1980). Temporal properties of binocular mechanisms in the human visual system, *Exp Brain Res* **38**: 333–340.

Blake, R., Overton, R. and Lema-Stern, S. (1981a). Interocular transfer of visual aftereffects, *J Exp Psychol Hum Percept Perform* **7**: 367–81.

Blake, R., Sloane, M. and Fox, R. (1981b). Further developments in binocular summation, *Percept Psychophys* **30**: 266–76.

Blakemore, C. (1970). A new kind of stereoscopic vision, *Vision Res* **10**: 1181–99.

Blakemore, C. and Campbell, F. W. (1969). On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images, *J Physiol* **203**: 237–60.

Blakemore, C. and Nachmias, J. (1971). The orientation specificity of two visual after-effects, *J Physiol* **213**: 157–74.

Blakemore, C. and Tobin, E. A. (1972). Lateral inhibition between orientation detectors in the cat's visual cortex, *Exp Brain Res* **15**: 439–40.

Bonds, A. B. (1989). Role of inhibition in the specification of orientation selectivity of cells in the cat striate cortex, *Vis Neurosci* **2**: 41–55.

Bonin, V., Mante, V. and Carandini, M. (2004). Nonlinear processing in LGN neurons, *in* S. Thrun, L. Saul and B. Schölkopf (eds), *Advances in Neural Information Processing Systems 16*, MIT Press, Cambridge MA.

Bonin, V., Mante, V. and Carandini, M. (2005). The suppressive field of neurons in lateral geniculate nucleus, *J Neurosci* **25**: 10844–56.

Boynton, G. M., Demb, J. B., Glover, G. H. and Heeger, D. J. (1999). Neuronal basis of contrast discrimination, *Vision Res* **39**: 257–69.

Bradley, A. and Freeman, R. D. (1981). Contrast sensitivity in anisometropic amblyopia, *Invest Ophthalmol Vis Sci* **21**: 467–76.

Brainard, D. H. (1997). The psychophysics toolbox, *Spat Vis* **10**: 433–6.

Buchert, M., Greenlee, M. W., Rutschmann, R. M., Kraemer, F. M., Luo, F. and Hennig, J. (2002). Functional magnetic resonance imaging evidence for binocular interactions in human visual cortex, *Exp Brain Res* **145**: 334–9.

Burton, G. J. (1981). Contrast discrimination by the human visual system, *Biol Cybern* **40**: 27–38.

Campbell, F. and Green, D. G. (1965). Monocular versus binocular visual acuity, *Nature* **208**: 191–192.

Campbell, F. W. and Kulikowski, J. J. (1966). Orientational selectivity of the human visual system, *J Physiol* **187**: 437–45.

Campbell, F. W. and Robson, J. G. (1968). Application of fourier analysis to the visibility of gratings, *J Physiol* **197**: 551–66.

Campbell, F. W., Cleland, B. G., Cooper, G. and Enroth-Cugell, C. (1968). The angular selectivity of visual cortical cells to moving gratings., *J Physiol* **198**: 237–250.

Campbell, F. W., Cooper, G. F. and Enroth-Cugell, C. (1969). The spatial selectivity of the visual cells of the cat., *J Physiol* **203**: 223–35.

Cannon, M. W. (1985). Perceived contrast in the fovea and periphery, *J Opt Soc Am A* **2**: 1760–8.

Carandini, M., Heeger, D. J. and Senn, W. (2002). A synaptic explanation of suppression in visual cortex, *J Neurosci* **22**: 10053–10065.

Cass, J. R. and Spehar, B. (2005). Dynamics of cross- and iso-surround facilitation suggest distinct mechanisms, *Vision Res* **45**: 3060–73.

Challinor, K. L., Meese, T. S. and Summers, R. J. (2007). Surround suppression saturates, cross-orientation suppression does not, *Perception* **36(S)**: 38.

Chino, Y. M., Cheng, H., Smith, E. L. r., Garraghty, P. E., Roe, A. W. and Sur, M. (1994). Early discordant binocular vision disrupts signal transfer in the lateral geniculate nucleus, *Proc Natl Acad Sci U S A* **91**: 6938–42.

Chirimuuta, M. and Tolhurst, D. J. (2005a). Accuracy of identification of grating contrast by human observers: Bayesian models of v1 contrast processing show correspondence between discrimination and identification performance, *Vision Res* **45**: 2960–71.

Chirimuuta, M. and Tolhurst, D. J. (2005b). Does a bayesian model of v1 contrast coding offer a neurophysiological account of human contrast discrimination?, *Vision Res* **45**: 2943–59.

Ciuffreda, K. J. and Fisher, S. K. (1987). Impairment of contrast discrimination in amblyopic eyes, *Ophthalmic Physiol Opt* **7**: 461–7.

Ciuffreda, K., Levi, D. M. and Selenow, A. (1990). *Amblyopia: Basic and Clinical Aspects*, Butterworth Heinemann.

Clatworthy, P. L., Chirimuuta, M., Lauritzen, J. S. and Tolhurst, D. J. (2003). Coding of the contrasts in natural images by populations of neurons in primary visual cortex (v1), *Vision Res* **43**: 1983–2001.

Cogan, A. I. (1987). Human binocular interaction: towards a neural model, *Vision Res* **27**: 2125–2139.

Cormack, L. K., Stevenson, S. B. and Landers, D. D. (1997). Interactions of spatial frequency and unequal monocular contrasts in stereopsis, *Perception* **26**: 1121–36.

Cornsweet, T. N. (1962). The staircase-method in psychophysics, *American Journal Of Psychology* **75**: 485–491.

Crewther, D. P., Jones, R., Munro, J., Price, T., Pulis, S. and Crewther, S. (2005). A neural network model of top-down rivalry, *in* D. Alais and R. Blake (eds), *Binocular Rivalry*, Bradford.

de Belsunce, S. and Sireteanu, R. (1991). The time course of interocular suppression in normal and amblyopic subjects, *Invest Ophthalmol Vis Sci* **32**: 2645–52.

De Valois, R. L., Albrecht, D. G. and Thorell, L. G. (1982a). Spatial frequency selectivity of cells in macaque visual cortex, *Vision Res* **22**: 545–59.

De Valois, R. L., Yund, E. W. and Hepler, N. (1982b). The orientation and direction selectivity of cells in macaque visual cortex, *Vision Res* **22**: 531–44.

de Weert, C. M. M. and Levelt, W. J. M. (1974). Binocular brightness combinations: Additive and nonadditive aspects, *Percept Psychophys* **15**: 551–562.

Dealy, R. S. and Tolhurst, D. J. (1974). Is spatial adaptation an after-effect of prolonged inhibition?, *J Physiol* **241**: 261–70.

DeAngelis, G. C., Robson, J. G., Ohzawa, I. and Freeman, R. D. (1992). Organization of suppression in receptive fields of neurons in cat visual cortex, *J Neurophysiol* **68**: 144–63.

Denny, N., Frumkes, T. E., Barris, M. C. and Eysteinsson, T. (1991). Tonic interocular suppression and binocular summation in human vision, *J Physiol* **437**: 449–60.

Derefeldt, G., Lennerstrand, G. and Lundh, B. (1979). Age variations in normal human contrast sensitivity, *Acta Ophthalmol* **57**: 679–90.

Derrington, A. M. and Lennie, P. (1984). Spatial and temporal contrast sensitivities of neurones in lateral geniculate nucleus of macaque, *J Physiol* **357**: 219–40.

DeValois, R. L. and DeValois, K. K. (1988). *Spatial vision*, Oxford University Press.

Ding, J. and Sperling, G. (2006a). Binocular combination: measurements and a model, *in* L. Harris and M. Jenkin (eds), *Computational vision in neural and machine systems*, Cambridge University Press.

Ding, J. and Sperling, G. (2006b). A gain-control theory of binocular combination, *Proc Natl Acad Sci U S A* **103**: 1141–6.

Eriksen, C. W. (1966). Independence of successive inputs and uncorrelated error in visual form perception, *J Exp Psychol* **72**: 26–35.

Fechner, G. T. (1860a). *Elemente der psychophysik*, Breitkopf und Härtel.

Fechner, G. T. (1860b). Uber einige verhältnisse des binokularen sehens, *Berichte sächs gesmte wissenschaft* **7**: 337 – 564.

Finney, D. J. (1971). *Probit Analysis*, Cambridge University Press.

Foley, J. M. (1994). Human luminance pattern-vision mechanisms: masking experiments require a new model, *J Opt Soc Am A Opt Image Sci Vis* **11**: 1710–1719.

Foley, J. M. and Boynton, G. M. (1994). A new model of human luminance pattern vision mechanisms: analysis of the effects of pattern orientation, spatial phase and temporal frequency, *in* T. B. Lawton (ed.), *Computational vision based on neurobiology, Proc. SPIE*, Vol. 2054, pp. 32–42.

Foley, J. M. and Chen, C. C. (1997). Analysis of the effect of pattern adaptation on pattern pedestal effects: a two-process model, *Vision Res* **37**: 2779–2788.

Foley, J. M. and Legge, G. E. (1981). Contrast detection and near-threshold discrimination in human vision, *Vision Res* **21**: 1041–1053.

Freeman, A. W. (2005). Multistage model for binocular rivalry, *J Neurophysiol* **94**: 4412–20.

Freeman, T. C., Durand, S., Kiper, D. C. and Carandini, M. (2002). Suppression without inhibition in visual cortex, *Neuron* **35**: 759–771.

Gagnon, R. W. C. and Kline, D. W. (2003). Senescent effects on binocular summation for contrast sensitivity and spatial interval acuity, *Curr Eye Res* **27**: 315–21.

García-Pérez, M. A. and Alcalá-Quintana, R. (2007). The transducer model for contrast detection and discrimination: formal relations, implications, and an empirical test., *Spat Vis* **20**: 5–43.

Geisler, W. S. and Albrecht, D. G. (1992). Cortical neurons: isolation of contrast gain control, *Vision Res* **32**: 1409–1410.

Geisler, W. S. and Albrecht, D. G. (1997). Visual cortex neurons in monkeys and cats: detection, discrimination, and identification, *Vis Neurosci* **14**: 897–919.

Georgeson, M. (2004). Visual aftereffects: cortical neurons change their tune, *Curr Biol* **14**: R751–3.

Georgeson, M. A. and Georgeson, J. M. (1987). Facilitation and masking of briefly presented gratings: time-course and contrast dependence, *Vision Res* **27**: 369–79.

Georgeson, M. A. and Harris, M. G. (1984). Spatial selectivity of contrast adaptation: models and data, *Vision Res* **24**: 729–41.

Georgeson, M. A. and Meese, T. S. (2005). Binocular summation at contrast threshold: a new look, *Perception* **34(S)**: 141.

Georgeson, M. A. and Meese, T. S. (2006). Fixed or variable noise in contrast discrimination? the jury's still out..., *Vision Res* **46**: 4294–303.

Georgeson, M. A. and Meese, T. S. (2007). Binocular combination at threshold: temporal filtering and summation of signals in separate on and off channels, *Perception* **36(S)**: 60.

Georgeson, M. A. and Sullivan, G. D. (1975). Contrast constancy: deblurring in human vision by spatial frequency channels., *J Physiol* **252**: 627–56.

Georgeson, M. A., Meese, T. S. and Baker, D. H. (2005). Binocular summation, dichoptic masking and contrast gain control, *J Vis* **5**: 797.

Gilchrist, J. and McIver, C. (1985). Fechner's paradox in binocular contrast sensitivity, *Vision Res* **25**: 609–613.

Gilchrist, J. and Pardhan, S. (1987). Binocular contrast detection with unequal monocular illuminance, *Ophthalmic Physiol Opt* **7**: 373–7.

Graham, N. and Nachmias, J. (1971). Detection of grating patterns containing two spatial frequencies: a comparison of single-channel and multiple-channels models., *Vision Res* **11**: 251–9.

Graham, N. V. S. (1989). *Visual Pattern Analyzers*, Oxford University Press.

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

Green, M. and Blake, R. (1981). Phase effects in monoptic and dichoptic temporal integration: flicker and motion detection, *Vision Res* **21**: 365–72.

Gregson, R. A. M. (1989). A nonlinear systems approach to fechner's paradox, *Biol Cybern* **61**: 129–138.

Hallett, P. E. (1969a). Quantum efficiency and false positive rate, *J Physiol* **202**: 421–36.

Hallett, P. E. (1969b). The variations in visual threshold measurement, *J Physiol* **202**: 403–19.

Halpern, D. L. and Blake, R. R. (1988). How contrast affects stereoacuity, *Perception* **17**: 483–95.

Halpern, D. L., Patterson, R. and Blake, R. (1987). Are stereoacuity and binocular rivalry related?, *Am J Optom Physiol Opt* **64**: 41–4.

Hammett, S. T., Champion, R. A., Morland, A. B. and Thompson, P. G. (2005). A ratio model of perceived speed in the human visual system, *Proc Biol Sci* **272**: 2351–6.

Hammett, S. T., Georgeson, M. A., Bedingham, S. and Barbieri-Hesse, G. S. (2003). Motion sharpening and contrast: gain control precedes compressive non-linearity?, *Vision Res* **43**: 1187–99.

Harrad, R. A. and Hess, R. F. (1992). Binocular integration of contrast information in amblyopia, *Vision Res* **32**: 2135–50.

Harris, M. G. (1980). Velocity specificity of the flicker to pattern sensitivity ratio in human vision, *Vision Res* **20**: 687–91.

Harwerth, R. S., Smith III, E. L. and Levi, D. M. (1980). Suprathreshold binocular interactions for grating patterns, *Percept Psychophys* **27**: 43–50.

Heeger, D. J. (1992). Normalization of cell responses in cat striate cortex, *Vis Neurosci* **9**: 181–197.

Henning, G. B. and Wichmann, F. A. (2007). Some observations on the pedestal effect, *J Vis* **7**: 1 – 15.

Heravian-Shandiz, J., Douthwaite, W. A. and Jenkins, T. C. (1991). Binocular interaction with neutral density filters as measured by the visual evoked response, *Optom Vis Sci* **68**: 801–806.

Hess, R. F. (1979). Contrast sensitivity assessment of functional amblyopia in humans, *Trans Ophthalmol Soc U K* **99**: 391–7.

Hess, R. F. (1983). Contrast-coding in amblyopia. ii. on the physiological basis of contrast recruitment, *Proc R Soc Lond B Biol Sci* **217**: 331–40.

Hess, R. F. (1996). Is amblyopia an impediment to binocular function?, *Eye* **10**: 245–9.

Hess, R. F. and Field, D. J. (1994). Is the spatial deficit in strabismic amblyopia due to loss of cells or an uncalibrated disarray of cells?, *Vision Res* **34**: 3397–406.

Hess, R. F. and Howell, E. R. (1977). The threshold contrast sensitivity function in strabismic amblyopia: evidence for a two type classification, *Vision Res* **17**: 1049–55.

Hess, R. F., Bradley, A. and Piotrowski, L. (1983). Contrast-coding in amblyopia. i. differences in the neural basis of human amblyopia, *Proc R Soc Lond B Biol Sci* **217**: 309–30.

Holmes, D. J. (2003). *Pooling and suppression in human spatial vision*, PhD thesis, Aston University.

Holmes, D. J. and Meese, T. S. (2004). Grating and plaid masks indicate linear summation in a contrast gain pool, *J Vis* **4**: 1080–1089.

Holopigian, K., Blake, R. and Greenwald, M. J. (1986). Selective losses in binocular vision in anisometropic amblyopes, *Vision Res* **26**: 621–30.

Holopigian, K., Blake, R. and Greenwald, M. J. (1988). Clinical suppression and amblyopia, *Invest Ophthalmol Vis Sci* **29**: 444–51.

Home, R. (1978). Binocular summation: a study of contrast sensitivity, visual acuity and recognition, *Vision Res* **18**: 579–85.

Hood, A. S. and Morrison, J. D. (2002). The dependence of binocular contrast sensitivities on binocular single vision in normal and amblyopic human subjects, *J Physiol* **540**: 607–22.

Howard, I. P. (2002). *Seeing in depth, vol I*, I Porteous, Toronto.

Huang, C., Tao, L., Zhou, Y. and Lu, Z.-L. (2007). Treated amblyopes remain deficient in spatial vision: A contrast sensitivity and external noise study, *Vision Res* **47**: 22–34.

Huang, P.-C. (2007). *Cortical properties of collinear facilitation*, PhD thesis, McGill University.

Huang, P.-C., Hess, R. F. and Dakin, S. C. (2006). Flank facilitation and contour integration: different sites, *Vision Res* **46**: 3699–706.

Hubel, D. H. (1959). Single unit activity in striate cortex of unrestrained cats, *J Physiol* **147**: 226–238.

Hubel, D. H. and Wiesel, T. N. (1959). Receptive fields of single neurones in the cat's striate cortex, *J Physiol* **148**: 574–91.

Hubel, D. H. and Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex., *J Physiol* **160**: 106–54.

Hubel, D. H. and Wiesel, T. N. (1965). Binocular interaction in striate cortex of kittens reared with artificial squint, *J Neurophysiol* **28**: 1041–59.

Ikeda, H. and Wright, M. J. (1975). Retinotopic distribution, visual latency and orientation tuning of "sustained" and "transient" cortical neurons in area 17 of the cat, *Exp Brain Res* **22**: 385–398.

Ikeda, H. and Wright, M. J. (1976). Properties of lgn cells in kittens reared with convergent squint: a neurophysiological demonstration of amblyopia, *Exp Brain Res* **25**: 63–77.

Kaernbach, C. (2001). Slope bias of psychometric functions derived from adaptive data, *Percept Psychophys* **63**: 1389–98.

Kalarickal, G. J. and Marshall, J. A. (2000). Neural model of temporal and stochastic properties of binocular rivalry, *Neurocomputing* **32-33**: 843–853.

Kertesz, A. E. and Jones, R. W. (1970). Human cyclofusional response, *Vision Res* **10**: 891–6.

Kiper, D. C. and Kiorpes, L. (1994). Suprathreshold contrast sensitivity in experimentally strabismic monkeys, *Vision Res* **34**: 1575–83.

Klein, S. A. (2006). Separating transducer non-linearities and multiplicative noise in contrast discrimination., *Vision Res* **46**: 4279–93.

Kontsevich, L. L., Chen, C.-C. and Tyler, C. W. (2002). Separating the effects of response nonlinearity and internal noise psychophysically, *Vision Res* **42**: 1771–84.

Kulikowski, J. J., Abadi, R. and King-Smith, P. E. (1973). Orientational selectivity of grating and line detectors in human vision, *Vision Res* **13**: 1479–86.

Langley, K. (2002). A parametric account of contrast adaptation on contrast perception, *Spat Vis* **16**: 77–93.

Lauritzen, J. S. and Tolhurst, D. J. (2005). Contrast constancy in natural scenes in shadow or direct light: A proposed role for contrast-normalisation (non-specific suppression) in visual cortex, *Network* **16**: 151–73.

Legge, G. E. (1978). Sustained and transient mechanisms in human vision: temporal and spatial properties, *Vision Res* **18**: 69–81.

Legge, G. E. (1979). Spatial frequency masking in human vision: binocular interactions, *J Opt Soc Am* **69**: 838–847.

Legge, G. E. (1984a). Binocular contrast summation–i. detection and discrimination., *Vision Res* **24**: 373–383.

Legge, G. E. (1984b). Binocular contrast summation–ii. quadratic summation., *Vision Res* **24**: 385–394.

Legge, G. E. and Foley, J. M. (1980). Contrast masking in human vision, *J Opt Soc Am* **70**: 1458–1471.

Legge, G. E. and Gu, Y. C. (1989). Stereopsis and contrast, *Vision Res* **29**: 989–1004.

Legge, G. E. and Rubin, G. S. (1981). Binocular interactions in suprathreshold contrast perception, *Percept Psychophys* **30**: 49–61.

Legge, G. E., Kersten, D. and Burgess, A. E. (1987). Contrast discrimination in noise, *J Opt Soc Am A* **4**: 391–404.

Lehky, S. R. (1983). A model of binocular brightness and binaural loudness perception in humans with general applications to nonlinear summation of sensory inputs, *Biol Cybern* **49**: 89–97.

Lehky, S. R. (1985). Temporal properties of visual channels measured by masking, *J Opt Soc Am A* **2**: 1260–72.

Lehky, S. R. (1988). An astable multivibrator model of binocular rivalry, *Perception* **17**: 215–28.

Lema, S. A. and Blake, R. (1977). Binocular summation in normal and stereoblind humans, *Vision Res* **17**: 691–695.

Lennie, P. and Movshon, J. A. (2005). Coding of color and form in the geniculostriate visual pathway (invited review), *J Opt Soc Am A Opt Image Sci Vis* **22**: 2013–33.

Leonards, U. and Sireteanu, R. (1993). Interocular suppression in normal and amblyopic subjects: the effect of unilateral attenuation with neutral density filters, *Percept Psychophys* **54**: 65–74.

Levelt, W. J. M. (1965a). Binocular brightness averaging and contour information, *Br J Psychol* **56**: 1–13.

Levelt, W. J. M. (1965b). *On binocular rivalry*, Institute for Perception RVO-TNO.

Levi, D. M. and Klein, S. A. (2003). Noise provides some new signals about the spatial vision of amblyopes, *J Neurosci* **23**: 2522–6.

Levi, D. M., Harwerth, R. and Smith III, E. L. (1979). Humans deprived of normal binocular vision have binocular interactions tuned to size and orientation, *Science* **206**: 852–854.

Levi, D. M., Harwerth, R. S. and Smith III, E. L. (1980). Binocular interactions in normal and anomalous binocular vision, *Doc Ophthalmol* **49**: 303–24.

Levi, D. M., Klein, S. A. and Chen, I. (2007). The response of the amblyopic visual system to noise, *Vision Res* **47**: 2531–42.

Levi, D. M., Klein, S. A. and Yap, Y. L. (1987). Positional uncertainty in peripheral and amblyopic vision, *Vision Res* **27**: 581–97.

Levitt, H. (1971). Transformed up-down methods in psychoacoustics, *J Acoust Soc Am* **49**: Suppl 2:467–477.

Li, B., Peterson, M. R., Thompson, J. K., Duong, T. and Freeman, R. (2005). Cross-orientation suppression: monoptic and dichoptic mechanisms are different, *J Neurophysiol* **94**: 1645–1650.

Li, B., Thompson, J. K., Duong, T., Peterson, M. R. and Freeman, R. D. (2006). Origins of cross-orientation suppression in the visual cortex, *J Neurophysiol* **96**: 1755–64.

Liu, L., Tyler, C. W. and Schor, C. M. (1992). Failure of rivalry at low contrast: evidence of a suprathreshold binocular summation process, *Vision Res* **32**: 1471–1479.

Luntinen, O., Rovamo, J. and Nasanen, R. (1995). Modelling the increase of contrast sensitivity with grating area and exposure time, *Vision Res* **35**: 2339–46.

Lynch, J. J. r., Silveira, L. C., Perry, V. H. and Merigan, W. H. (1992). Visual effects of damage to p ganglion cells in macaques, *Vis Neurosci* **8**: 575–83.

Maattanen, L. M. and Koenderink, J. J. (1991). Contrast adaptation and contrast gain control, *Exp Brain Res* **87**: 205–12.

Maehara, G. and Goryo, K. (2005). Binocular, monocular and dichoptic pattern masking, *Optical Review* **12**: 76–82.

Maffei, L. and Fiorentini, A. (1973). The visual cortex as a spatial frequency analyser, *Vision Res* **13**: 1255–67.

Manahilov, V., Gordon, G., Calvert, J. and Simpson, W. A. (2007). Subtractive suppression underlies contrast processing of visual stimuli, *Perception* **36(S)**: 307.

Marr, D. (1983). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*, W. H. Freeman.

Marshman, W. E., Dawson, E., Neveu, M. M., Morgan, M. J. and Sloper, J. J. (2001). Increased binocular enhancement of contrast sensitivity and reduced stereoacuity in duane syndrome, *Invest Ophthalmol Vis Sci* **42**: 2821–5.

McIlhagga, W. and Peterson, R. (2006). Sinusoid = light bar + dark bar?, *Vision Res* **46**: 1934–45.

Medina, J. M. and Mullen, K. T. (2007). Colour-luminance interactions in binocular summation, *Vision Res* **47**: 1120–8.

Medina, J. M., Meese, T. S. and Mullen, K. T. (2007). Cross-orientation masking in the red-green isoluminant and luminance systems, *J Vis* **7**: 257a.

Meese, T. (2004). Area summation and masking, *J Vis* **4**: 930–943.

Meese, T. and Holmes, D. J. (2002). Adaptation and gain pool summation: alternative models and masking data, *Vision Res* **42**: 1113–1125.

Meese, T. S. and Hess, R. F. (2004). Low spatial frequencies are suppressively masked across spatial scale, orientation, field position, and eye of origin, *J Vis* **4**: 843–859.

Meese, T. S. and Hess, R. F. (2005). Interocular suppression is gated by interocular feature matching, *Vision Res* **45**: 9–15.

Meese, T. S. and Holmes, D. J. (2003). Orientation-masking: suppression and mechanism bandwidths, *Perception* **32**: 388.

Meese, T. S. and Holmes, D. J. (2006). Cross-orientation suppression is proportional to the square-root of speed for flickering gabor stimuli, *J Vis* **6**: 200.

Meese, T. S. and Holmes, D. J. (2007). Spatial and temporal dependencies of cross-orientation suppression in human vision., *Proc Biol Sci* **274**: 127–36.

Meese, T. S. and Summers, R. J. (2007). Area summation in human vision at and above detection threshold., *Proc Biol Sci* **274**: 2891–900.

Meese, T. S., Georgeson, M. A. and Baker, D. H. (2005a). Interocular masking and summation indicate two stages of divisive contrast gain control, *Perception* **34**(S): 42–43.

Meese, T. S., Georgeson, M. A. and Baker, D. H. (2006). Binocular contrast vision at and above threshold, *J Vis* **6**: 1224–1243.

Meese, T. S., Hess, R. F. and Williams, C. B. (2005b). Size matters, but not for everyone: individual differences for contrast discrimination, *J Vis* **5**: 928–47.

Meese, T. S., Holmes, D. J. and Challinor, K. L. (2007a). Remote facilitation in the fourier domain, *Vision Res* **47**: 1112–9.

Meese, T. S., Summers, R. J., Holmes, D. J. and Wallis, S. A. (2007b). Contextual modulation involves suppression and facilitation from the centre and the surround, *Journal of Vision* **7**: 1 – 21.

Merigan, W. H. and Eskin, T. A. (1986). Spatio-temporal vision of macaques with severe loss of p beta retinal ganglion cells, *Vision Res* **26**: 1751–61.

Merigan, W. H., Byrne, C. E. and Maunsell, J. H. (1991a). Does primate motion perception depend on the magnocellular pathway?, *J Neurosci* **11**: 3422–9.

Merigan, W. H., Katz, L. M. and Maunsell, J. H. (1991b). The effects of parvocellular lateral geniculate lesions on the acuity and contrast sensitivity of macaque monkeys, *J Neurosci* **11**: 994–1001.

Moore, B. C. J. and Glasberg, B. R. (1996). A revision of Zwicker's loudness model, *Acustica* **82**: 335–345.

Morrone, M. C., Burr, D. C. and Maffei, L. (1982). Functional implications of cross-orientation inhibition of cortical visual cells. i. neurophysiological evidence, *Proc R Soc Lond B Biol Sci* **216**: 335–54.

Movshon, J. A. and Lennie, P. (1979). Pattern-selective adaptation in visual cortical neurones, *Nature* **278**: 850–2.

Movshon, J. A., Thompson, I. D. and Tolhurst, D. J. (1978). Spatial and temporal contrast sensitivity of neurones in areas 17 and 18 of the cat's visual cortex, *J Physiol* **283**: 101–20.

Nachmias, J. (1967). Effect of exposure duration on the visual contrast sensitivity with square-wave gratings, *J Opt Soc Am* **57**: 421–427.

Nachmias, J. (1981). On the psychometric function for contrast detection, *Vision Res* **21**: 215–23.

Nachmias, J. (2006). The role of virtual standards in visual discrimination, *Vision Res* **46**: 2456–64.

Nachmias, J. and Sansbury, R. V. (1974). Letter: Grating contrast: discrimination may be better than detection, *Vision Res* **14**: 1039–1042.

Naka, K. I. and Rushton, W. A. (1966). S-potentials from colour units in the retina of fish (cyprinidae), *J Physiol* **185**: 536–55.

Nelder, J. A. and Mead, R. (1965). A simplex method for function minimization, *Computer Journal* **7**: 308–313.

Nelson, S. B. (1991). Temporal interactions in the cat visual system. i. orientation-selective suppression in the visual cortex, *J Neurosci* **11**: 344–56.

Ohzawa, I. and Freeman, R. D. (1994). Monocular and binocular mechanisms of contrast gain control, *Computational vision based on neurobiology* **2054**: 43–51.

Ohzawa, I., Sclar, G. and Freeman, R. D. (1982). Contrast gain control in the cat visual cortex, *Nature* **298**: 266–8.

Ohzawa, I., Sclar, G. and Freeman, R. D. (1985). Contrast gain control in the cat's visual system, *J Neurophysiol* **54**: 651–67.

Olzak, L. A. and Thomas, J. P. (2003). Dual nonlinearities regulate contrast sensitivity in pattern discrimination tasks, *Vision Res* **43**: 1433–1442.

Pardhan, S. (1996). A comparison of binocular summation in young and older patients, *Curr Eye Res* **15**: 315–9.

Pardhan, S. and Gilchrist, J. (1992). Binocular contrast summation and inhibition in amblyopia. the influence of the interocular difference on binocular contrast sensitivity, *Doc Ophthalmol* **82**: 239–248.

Pardhan, S. and Rose, D. (1999). Binocular and monocular detection of gabor patches in binocular two-dimensional noise, *Perception* **28**: 203–15.

Pardhan, S. and Whitaker, A. (2000). Binocular summation in the fovea and peripheral field of anisometropic amblyopes, *Curr Eye Res* **20**: 35–44.

Peirce, J. W. (2007). The potential importance of saturating and supersaturating contrast response functions in visual cortex, *J Vis* **7**: 13.

Pelli, D. G. (1981). *The effects of visual noise*, PhD thesis, Cambridge University.

Pelli, D. G. (1985). Uncertainty explains many aspects of visual contrast detection and discrimination, *J Opt Soc Am A* **2**: 1508–32.

Pelli, D. G. (1987). On the relation between summation and facilitation, *Vision Res* **27**: 119–23.

Pelli, D. G. (1997). The videotoolbox software for visual psychophysics: transforming numbers into movies, *Spat Vis* **10**: 437–42.

Petrov, Y. and McKee, S. P. (2006). The effect of spatial configuration on surround suppression of contrast sensitivity, *J Vis* **6**: 224–38.

Petrov, Y., Carandini, M. and McKee, S. (2005). Two distinct mechanisms of suppression in human vision, *J Neurosci* **25**: 8704–7.

Petrov, Y., Verghese, P. and McKee, S. P. (2006). Collinear facilitation is largely uncertainty reduction, *J Vis* **6**: 170–8.

Pettigrew, J. D. and Miller, S. M. (1998). A 'sticky' interhemispheric switch in bipolar disorder?, *Proc Biol Sci* **265**: 2141–8.

Phillips, G. and Wilson, H. (1984). Orientation bandwidths of spatial mechanisms measured by masking, *J Opt Soc Am A* **1**: 226–232.

Pirenne, M. H. (1943). Binocular and uniocular threshold of vision, *Nature* **152**: 698–699.

Priebe, N. J. and Ferster, D. (2006). Mechanisms underlying cross-orientation suppression in cat visual cortex, *Nat Neurosci* **9**: 552–61.

Quick, R. F. (1974). A vector-magnitude model of contrast detection, *Kybernetik* **16**: 65–67.

Rainville, S. J. M., Makous, W. L. and Scott-Samuel, N. E. (2005). Opponent-motion mechanisms are self-normalizing, *Vision Res* **45**: 1115–27.

Robson, J. G. and Graham, N. (1981). Probability summation and regional variation in contrast sensitivity across the visual field, *Vision Res* **21**: 409–18.

Rose, A. (1948). The sensitivity performance of the human eye on an absolute scale, *J Opt Soc Am* **38**: 196 – 208.

Rose, D. (1978). Monocular versus binocular contrast thresholds for movement and pattern, *Perception* **7**: 195–200.

Rose, D. (1980). The binocular: monocular sensitivity ratio for movement detection varies with temporal frequency, *Perception* **9**: 577–80.

Rose, D., Blake, R. and Halpern, D. L. (1988). Disparity range for binocular summation, *Invest Ophthalmol Vis Sci* **29**: 283–90.

Ross, J. and Speed, H. D. (1996). Perceived contrast following adaptation to gratings of different orientations, *Vision Res* **36**: 1811–1818.

Ross, J. E., Clarke, D. D. and Bron, A. J. (1985). Effect of age on contrast sensitivity function: uniocular and binocular findings, *Br J Ophthalmol* **69**: 51–6.

Rutstein, R. P. (2005). Contemporary issues in amblyopia treatment, *Optometry* **76**: 570–8.

Sachs, M. B., Nachmias, J. and Robson, J. G. (1971). Spatial-frequency channels in human vision, *J Opt Soc Am* **61**: 1176–86.

Sasaki, H., Todorokihara, M., Ishida, T., Miyachi, J., Kitamura, T. and Aoki, R. (2006). Effect of noise on the contrast detection threshold in visual perception, *Neurosci Lett* **408**: 94–7.

Schade, O. H. S. (1956). Optical and photoelectric analog of the eye., *J Opt Soc Am* **46**: 721–39.

Schor, C. and Heckmann, T. (1989). Interocular differences in contrast and spatial frequency: effects on stereopsis and fusion, *Vision Res* **29**: 837–47.

Sengpiel, F. and Blakemore, C. (1996). The neural basis of suppression and amblyopia in strabismus, *Eye* **10**: 250–8.

Sengpiel, F. and Vorobyov, V. (2005). Intracortical origins of interocular suppression in the visual cortex, *J Neurosci* **25**: 6394–6400.

Sengpiel, F., Blakemore, C. and Harrad, R. (1995). Interocular suppression in the primary visual cortex: a possible neural basis of binocular rivalry, *Vision Res* **35**: 179–195.

Sengpiel, F., Freeman, T. C., Bonhoeffer, T. and Blakemore, C. (2001). On the relationship between interocular suppression in the primary visual cortex and binocular rivalry, *Brain and Mind* **2**: 39–54.

Sengpiel, F., Jirmann, K.-U., Vorobyov, V. and Eysel, U. T. (2006). Strabismic suppression is mediated by inhibitory interactions in the primary visual cortex, *Cereb Cortex* **16**: 1750–8.

Sherrington, C. S. (1904). On binocular flicker and the correlation of activity of corresponding retinal points, *Br J Psychol* **1**: 29 – 60.

Simmons, D. R. (2005). The binocular combination of chromatic contrast, *Perception* **34**: 1035–42.

Simmons, D. R. and Kingdom, F. A. (1998). On the binocular summation of chromatic contrast, *Vision Res* **38**: 1063–71.

Sireteanu, R. and Singer, W. (1980). The "vertical effect" in human squint amblyopia, *Exp Brain Res* **40**: 354–7.

Sireteanu, R., Fronius, M. and Singer, W. (1981). Binocular interaction in the peripheral visual field of humans with strabismic and anisometropic amblyopia, *Vision Res* **21**: 1065–74.

Smith III, E. L., Harweth, R. S., Levi, D. M. and Boltz, R. L. (1982). Contrast increment thresholds of rhesus monkeys, *Vision Res* **22**: 1153–61.

Smith, M. A., Bair, W. and Movshon, J. A. (2006). Dynamics of suppression in macaque primary visual cortex, *J Neurosci* **26**: 4826–34.

Snowden, R. and Hammett, S. (1996). Spatial frequency adaptation: threshold elevation and perceived contrast, *Vision Res* **36**: 1797–1809.

Sperling, G. and Ding, J. (2006). An early gain-control mechanism in binocular combination, *Journal of Vision* **6**: 832a.

Stollenwerk, L. and Bode, M. (2003). Lateral neural model of binocular rivalry, *Neural Computation* **15**: 2863–2882.

Strasburger, H. (2001). Converting between measures of slope of the psychometric function, *Percept Psychophys* **63**: 1348–55.

Stromeyer, C. F. I. and Klein, S. (1975). Evidence against narrow-band spatial frequency channels in human vision: the detectability of frequency modulated gratings, *Vision Res* **15**: 899–910.

Summers, R. J. and Meese, T. S. (2007). The influence of fixation points on contrast detection of patches of grating, *Perception*.

Swanson, W. H., Wilson, H. and Giese, S. C. (1984). Contrast matching data predicted from contrast increment thresholds, *Vision Res* **24**: 63–75.

Tolhurst, D. and Heeger, D. J. (1997). Comparison of contrast-normalization and threshold models of the responses of simple cells in cat striate cortex, *Vis Neurosci* **14**: 293–309.

Tolhurst, D. J. (1975a). Reaction times in the detection of gratings by human observers: a probabilistic mechanism, *Vision Res* **15**: 1143–9.

Tolhurst, D. J. (1975b). Sustained and transient channels in human vision, *Vision Res* **15**: 1151–5.

Tolhurst, D. J. and Barfield, L. P. (1978). Interactions between spatial frequency channels, *Vision Res* **18**: 951–8.

Tolhurst, D. J. and Dean, A. F. (1987). Spatial summation by simple cells in the striate cortex of the cat, *Exp Brain Res* **66**: 607–20.

Tolhurst, D. J. and Thompson, I. D. (1981). On the variety of spatial frequency selectivities shown by neurons in area 17 of the cat, *Proc R Soc Lond B Biol Sci* **213**: 183–99.

Tolhurst, D. J., Dean, A. F. and Thompson, I. D. (1981a). Preferred direction of movement as an element in the organization of cat visual cortex, *Exp Brain Res* **44**: 340–2.

Tolhurst, D. J., Movshon, J. A. and Thompson, I. D. (1981b). The dependence of response amplitude and variance of cat visual cortical neurones on stimulus contrast, *Exp Brain Res* **41**: 414–9.

Tolkien, J. R. R. (1954). *The fellowship of the ring*, Harper Collins.

Truchard, A. M., Ohzawa, I. and Freeman, R. D. (2000). Contrast gain control in the visual cortex: monocular versus binocular mechanisms, *J Neurosci* **20**: 3017–32.

Tyler, C. W. and Chen, C. C. (2000). Signal detection theory in the 2afc paradigm: attention, channel uncertainty and probability summation, *Vision Res* **40**: 3121–3144.

Valberg, A. and Fosse, P. (2002). Binocular contrast inhibition in subjects with age-related macular degeneration, *J Opt Soc Am A Opt Image Sci Vis* **19**: 223–228.

Vedamurthy, I., Suttle, C. M., Alexander, J. and Asper, L. J. (2007). Interocular interactions during acuity measurement in children and adults, and in adults with amblyopia, *Vision Res* **47**: 179–88.

von Noorden, G. K. and Crawford, M. L. (1992). The lateral geniculate nucleus in human strabismic amblyopia, *Invest Ophthalmol Vis Sci* **33**: 2729–32.

Watson, A. B. (2000). Visual detection of spatial contrast patterns: evaluation of five simple models, *Opt Express* **6**: 12–33.

Watson, A. B. and Ahumada, A. J. J. (2005). A standard model for foveal detection of spatial contrast, *J Vis* **5**: 717–40.

Watson, A. B. and Solomon, J. A. (1997). Model of visual contrast gain control and pattern masking, *J Opt Soc Am A Opt Image Sci Vis* **14**: 2379–91.

Watson, A. B., Borthwick, R. and Taylor, M. (1997). Image quality and entropy masking, *SPIE Proceedings* **3016**: 1–11.

Webb, B. S., Dhruv, N. T., Solomon, S. G., Tailby, C. and Lennie, P. (2005). Early and late mechanisms of surround suppression in striate cortex of macaque, *J Neurosci* **25**: 11666–75.

Wertheimer, M. (1953). An investigation of the randomness of threshold measurements, *J Exp Psychol* **45**: 294–303.

Westendorf, D. H. and Fox, R. (1974). Binocular detection of positive and negative flashes, *Percept Psychophys* **15**: 61 – 65.

Westendorf, D. H., Langston, A., Chambers, D. and Allegretti, C. (1978). Binocular detection by normal and stereoblind observers, *Percept Psychophys* **24**: 209–214.

Wichmann, F. A. (1999). *Some aspects of modelling human spatial vision: contrast discrimination*, PhD thesis, Oxford University.

Wiesenfeld, K. and Moss, F. (1995). Stochastic resonance and the benefits of noise: from ice ages to crayfish and squids, *Nature* **373**: 33–36.

Wilson, H. (1980). A transducer function for threshold and suprathreshold human vision, *Biol Cybern* **38**: 171–178.

Wilson, H. (1999). *Spikes, Decisions, and Actions: The Dynamical Foundations of Neuroscience*, Oxford University Press.

Wilson, H. (2003). Computational evidence for a rivalry hierarchy in vision, *Proc Natl Acad Sci USA* **100**: 14499–14503.

Wilson, H. and Humanski, R. (1993). Spatial frequency adaptation and contrast gain control, *Vision Res* **33**: 1133–1149.

Wilson, H. R. and Bergen, J. R. (1979). A four mechanism model for threshold spatial vision, *Vision Res* **19**: 19–32.

Wilson, H. R., McFarlane, D. K. and Phillips, G. C. (1983). Spatial frequency tuning of orientation selective units estimated by oblique masking, *Vision Res* **23**: 873–82.

Wolfe, J. M. (1983). Influence of spatial frequency, luminance, and duration on binocular rivalry and abnormal fusion of briefly presented dichoptic stimuli, *Perception* **12**: 447–456.

Yang, J. and Stevenson, S. B. (1999). Post-retinal processing of background luminance, *Vision Res* **39**: 4045–51.

Yang, J., Qi, X. and Makous, W. (1995). Zero frequency masking and a model of contrast sensitivity, *Vision Res* **35**: 1965–78.

Yu, C., Klein, S. and Levi, D. M. (2002). Facilitation of contrast detection by cross-oriented surround stimuli and its psychophysical mechanisms, *J Vis* **2**: 243–255.

Appendix A

Stereoscope calibration

A.1 Introduction

The mirror stereoscope was invented in the early 19th century by Sir Charles Wheatstone. It allows presentation of different images to the two eyes using appropriately angled mirrors. Despite the wide usage of stereoscopes for research on binocular combination, stereopsis, and binocular rivalry, there are no published guidelines on how best to calibrate a stereoscope for use in psychophysical experiments.

Some stereoscopes require that the observer fuses two monocular images by making a horizontal vergence movement of the eyes. Such devices do not require calibration, but rely on the observer to achieve stereo fusion, and also to indicate when it has been lost. A more versatile device, such as that used throughout this thesis, can be calibrated so that vergence is automatic. This is achieved by the correct angling of mirrors, and thus requires individual calibration for each observer, to account for differences in interocular separation.

A.2 Design and theory

The basic layout of the stereoscope is shown in Figure A.1[1]. The eyes are placed in front of two mirrors, angled obliquely, which reflect images from a further two oblique mirrors, which reflect stimuli presented on the display. The distance between the inner and outer mirrors defines the lateral displacement of the images shown on the display. In the diagram, this is 6cm.

When properly calibrated, the eyes converge on a central 'virtual object'. Although

[1]The stereoscope under discussion was constructed from standard optical bench components, according to a design on Randolph Blake's website: http://www.psy.vanderbilt.edu/faculty/blake/Stereoscope/stereoscope.html.

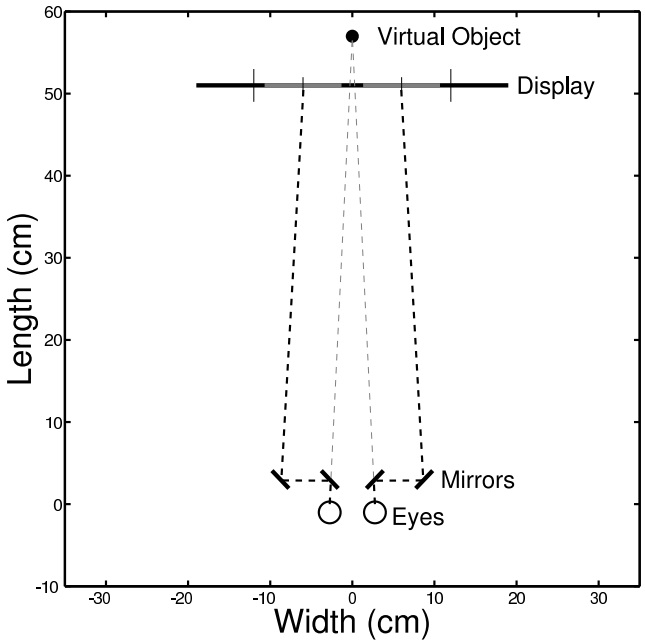

Figure A.1: Schematic diagram of the stereoscope set up. Images shown in the two halves of the display are horizontally displaced by pairs of mirrors so that they appear central and spatially coincident when viewed binocularly. The pale dashed lines show the eyes converging on the 'virtual object', which appears more distant than the display. The darker dashed lines show the true path which light takes from the display to the eyes.

the physical distance between eye and screen here is 51cm, the object appears to be 57cm away, as the light has travelled a further 6cm due to the mirrors. Thus, the optical viewing distance for the device is 57cm. Both the true and virtual angles of gaze are shown by the dashed lines in Figure A.1.

If the eyes were aimed directly forward (no convergence), then the appropriate mirror angle would be 45°, and the images for fusion should be displaced from the centre by 6cm (mirror separation) plus half the interocular separation. However, given that the eyes are focussed on a central 'virtual image', the appropriate mirror angle will change slightly. These angles will be different for different interocular separations, which vary across observers, and are generally between 5 and 7cm. Although the appropriate adjustment can be calculated using basic geometry (usually $\sim 1.5°$), it is very difficult indeed to accurately measure the angles of the mirrors. Instead, a subjective calibration procedure was developed to calibrate the stereoscope.

A.3 Calibration procedure

First the observer's interocular separation is measured using a standard technique familiar to optometrists. A ruler is held across the observer's forehead, spanning the distance between

the eyes. The observer fixates the left eye of the experimenter, and the location of the right pupil on the ruler is noted. The observer then fixates the right eye of the experimenter, and the location of the left pupil is read from the ruler. This procedure provides a relatively accurate estimate of the interpupillary distance when the observer is looking directly ahead. This ranges between 5 and 8cm for humans.

The central pair of mirrors are adjusted to this width, taking care that they are equidistant from the centre of the stereoscope, which in turn must be aligned with the centre of the display. The outer mirrors are then moved 6cm away from the inner mirrors, and all mirrors are secured in place using screws.

Each mirror has several degrees of freedom (horizontal and vertical displacement and rotation around the vertical and horizontal axes). They are first set to $\pm 45°$, as shown in Figure A.1, and fixed in place. Small precision screws then allow for finer adjustment of both the horizontal and vertical alignment of the mirror. It is these which are used to accurately calibrate the stereoscope.

The observer first views a blank screen of mean luminance through the stereoscope, with the rest of the room in darkness to prevent any unwanted reflections. In each central mirror, a luminous disc is apparent, caused by the reflection of mean luminance in the outer mirrors. The observer adjusts the mirrors so that these discs are roughly equidistant from the vertical midline (i.e. the observer's own nose).

The calibration image shown in Figure A.2 is then displayed, and the device calibrated separately for each eye, keeping the other eye closed or patched whilst doing so. For each eye, the large black circle should appear in the centre of the disc of light reflected from the outer mirror. Horizontal and vertical adjustments are made to achieve this.

Precise horizontal and vertical alignment is then carried out. The basic technique is to compare the image seen in the mirror to the real image displayed on the screen. This may require slight vertical or horizontal movements of the head, in order that both the reflected and real images can be viewed at once. Vertical alignment makes use of the horizontal line running through the centre of the calibration image. The vertical tilt of the inner and outer mirrors is adjusted such that the horizontal line in the mirror is collinear with that displayed on the screen. Both should appear precisely horizontal, and be continuous. This is usually relatively straightforward.

Horizontal alignment makes use of the long vertical lines in the upper half of the image. For the right eye, the far right line in the centre of the circle should be seen in the mirror. The horizontal tilt of the mirrors is then adjusted so that this line is collinear with the rightmost line of the central pair, as displayed on the screen and viewed over the top of the mirror (there are two central lines because the virtual image appears further away from that

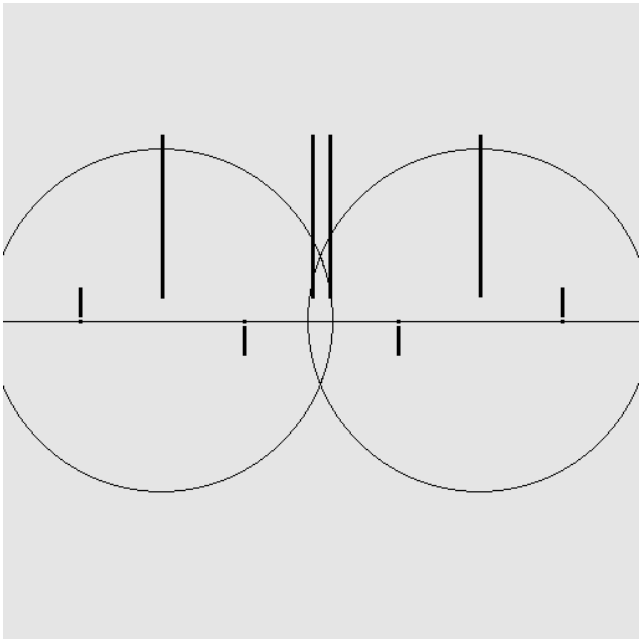

Figure A.2: Image developed for use in the calibration procedure. The background was set to mean luminance. The large circles are displaced laterally by an amount equal to the distance between inner and outer mirrors. For the setup described here, this was always 6cm. Horizontal and vertical lines aided adjustment of the stereoscope mirrors, as described in the text.

shown on the screen - see the points at which the pale dashed lines in Figure A.1 intersect the display). This adjustment may take a little time and practise, as both mirrors may need to be manipulated to produce a good match.

When horizontal and vertical alignment has been carried out for both eyes, a final check is to use the two fixation points and nonius line pairs along the central line. If the fixation points are fused and stable, and the nonius line pairs appear collinear, then calibration has been successful (emphasis is placed on horizontal alignment, as the stereoscope is designed to displace images horizontally). Images presented in the centres of the two circles will now be shown to the left and right eyes, but appear to occupy the same point in space.