SOME ASPECTS OF VISUAL DISCOMFORT

Louise O’Hare

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

2013

Full metadata for this item is available in
Research@StAndrews:FullText
at:
http://research-repository.st-andrews.ac.uk/

Please use this identifier to cite or link to this item:
http://hdl.handle.net/10023/3551

This item is protected by original copyright
Some Aspects of Visual Discomfort

Louise O’Hare
School of Psychology and Neuroscience
University of St Andrews

Supervised by:
Dr Paul B. Hibbard
and
Dr Dhanraj Vishwanath

Thesis submitted in partial fulfilment for the
degree of Doctor of Philosophy

September 2012
DECLARATION

I, Louise O’Hare, hereby certify that this thesis, which is approximately 77,245 words in length, has been written by me, that it is the record of work carried out by me and that it has not been submitted in any previous application for a higher degree.

I was admitted as a research student in September 2009 and as a candidate for the degree of doctor of philosophy in September 2012; the higher study for which this is a record was carried out in the University of St Andrews between 2009 and 2012.

Date........................................................

Signature of candidate......................................

2. Supervisor’s declaration:

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

Date........................................................

Signature of supervisor......................................

3. Permission for electronic publication:

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

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

Access to all of the printed copy but embargo part of the electronic publication of thesis for a period of two years as publication would preclude future publication.

Date........................................................

Signature of candidate......................................

Signature of supervisor....................................
ABSTRACT

Visual discomfort is the adverse sensations, such as headaches and eyestrain, encountered on viewing certain stimuli. These sensations can arise under certain viewing conditions, such as stereoscopic viewing and prolonged reading of text patterns. Also, discomfort can occur as a result of viewing stimuli with certain spatial properties, including stripes and filtered noise patterns of particular spatial frequency. This thesis is an exploration of the stimulus properties causing discomfort, within the framework of two theoretical explanations. Both of the explanations relate to the stimuli being difficult for the visual system to process. The first is concerned with discomfort being the result of inefficient neural processing. Neural activity requires energy to process information, and stimuli that demand a lot of energy to be processed might be uncomfortable. The second explanation revolves around uncomfortable stimuli not being effective in driving the accommodative (focussing) response. Accommodation relies on the stimulus as a cue to drive the response effectively - an uninformative cue might result in discomfort from an uncertain accommodative response. The following research investigates both these possibilities using a combination of psychophysical experimentation, questionnaire-based surveys on non-clinical populations, and computational modelling. The implications of the work for clinical populations are also discussed.
ACKNOWLEDGEMENTS

I would like to thank all the people who made this possible - there are so many who contributed in one way or another that I am only able to mention a small number individually. To single out a few, thanks to those of practical assistance, such as the BBSRC and Guarantors of Brain for funding, and the support staff of the University of St Andrews School of Psychology and Neuroscience. Thanks to all those who shared their expertise, in whatever form, whether training, tips, code, or other general advice. In particular, thanks to my collaborators, Harold Nefs, Tingting Zhang, Catherine Gordon, Caroline Foubister, the people of NCR, Alasdair Clarke, and of course the Vision lab members past and present. I would like to thank all my participants, the data would have been severely impoverished without those prepared to spend considerable amounts of time looking at unpleasant pictures. Thank you to Dhanraj, for being there should something have gone wrong. A very special thanks to Paul, for being a patient teacher, a diplomatic critic, an abundant source of useful pointers, in short, an all-round fantastic supervisor. Finally, thanks to my family and friends, especially Mum, Dad, Matt and Soph, and both Grandmas, for all the love, support, assistance and confidence that this would be possible. Thank you all so much for everything.
COLLABORATION STATEMENT

Chapter 5 - Depth of Field and Discomfort This was in collaboration with Harold Nefs and Tingting Zhang of the TU Delft. The experimental design was my own. I collected the data, and analysed it. The contribution of Harold Nefs and Tingting Zhang was in discussion of the results. My initial interpretation of results was later discussed with Harold Nefs and Tingting Zhang.

Chapter 6 - Effect of Uncomfortable Images on Visual Search This set of experiments was in collaboration with Alasdair Clarke. Alasdair Clarke was involved in discussions throughout all the experiments. Initial concept, experimental design and method of analysis were the result of discussions between Alasdair Clarke and myself. Stimuli were created, data were collected, analysed and initially interpreted by myself. Interpretation of results was discussed with Alasdair Clarke.

Chapter 7 - Effects of Visual Discomfort on Reading Experiment 1 was in collaboration with Catherine Gordon. Data for Experiment 1 were collected by Catherine Gordon. Experimental design, stimulus design, analysis and interpretation were done by myself. Catherine Gordon also assisted with the preparation of the stimuli for Experiment 1.

Chapter 8 - Effects of Contrast and Luminance on Visual Discomfort Experiment 1 was without any input from other parties. Experiment 2 and 3 were in collaboration with Caroline Foubister, and some staff at NCR. The staff of NCR loaned the high-bright display used in Experiment 3. Experimental design and data collection for Experiment 2 were done by Caroline Foubister for her undergraduate project work. Stimuli were also created by Caroline Foubister. Caroline Foubister analysed the data separately and submitted this separately for credit for her undergraduate degree. I separately analysed and interpreted the data for Experiment 2 with no input from Caroline Foubister. Experiment 3 used the same materials and design as Experiment 2 for ease of comparison between the two experiments. Data collection, analysis and interpretation were done by myself for Experiment 3.
## CONTENTS

Abstract iii

Acknowledgements iv

Collaboration Statement v

Table of Contents vi

List of Figures ix

1 Introduction 1
   1.1 Visual Discomfort ................................................. 1
   1.2 Natural Images ..................................................... 6
   1.3 Sparse Coding ...................................................... 8
   1.4 Accommodation Responses ......................................... 11
   1.5 Summary ........................................................... 12

2 Spatial Frequency and Visual Discomfort 15
   2.1 Introduction ......................................................... 15
   2.2 General Methods ..................................................... 20
   2.3 Experiment 1: Contrast Matching ................................. 22
   2.4 Experiment 2: Relative Discomfort
       Judgements ........................................................... 26
   2.5 Experiment 3: Contrast Matching for Centrally Presented Stimuli 32
   2.6 Experiment 4: Relative Discomfort
       Judgements for Centrally Presented
       Stimuli ................................................................. 33
   2.7 Experiment 5: Discomfort Rankings for Printed Stimuli ......... 36
   2.8 General Discussion .................................................. 39
## CONTENTS

3 Modelling Sparse Coding .......................... 43
  3.1 Introduction ........................................ 43
  3.2 The Model .......................................... 47
  3.3 Results ............................................. 60
  3.4 Discussion ......................................... 69

4 Blur and Visual Discomfort ..................... 79
  4.1 Introduction ........................................ 79
  4.2 General Methods .................................... 85
  4.3 Experiment 1: Discomfort Judgements of Simple Stimuli .... 87
  4.4 Experiment 2: Blur Judgements of Simple Stimuli ............ 93
  4.5 Experiment 3: Blurring Natural Images ..................... 95
  4.6 An Idealised Model of Defocus Blur ..................... 99
  4.7 Conclusion ......................................... 110

5 Depth-of-Field and Visual Discomfort ........ 114
  5.1 Introduction ........................................ 114
  5.2 Method .............................................. 118
  5.3 Results ............................................. 125
  5.4 Discussion ......................................... 130

6 Effect of Uncomfortable Images on Visual Search 134
  6.1 Introduction ........................................ 134
  6.2 Experiment 1: Filtered Noise Backgrounds ............... 137
  6.3 Experiment 2: Striped Backgrounds .................... 144
  6.4 Experiment 3: Target Orientation ..................... 150
  6.5 Conclusion ......................................... 151

7 Effects of Discomfort on Reading ............. 154
  7.1 Introduction ........................................ 154
  7.2 Experiment 1 ....................................... 162
  7.3 Experiment 2 ....................................... 172

8 Effects of Contrast and Luminance on Visual Discomfort 178
  8.1 Introduction ........................................ 178
  8.2 Experiment 1 ....................................... 182
  8.3 Experiment 2: CRT Monitor ......................... 186
  8.4 Experiment 3: High-Bright Display ................... 194
  8.5 General Discussion ................................ 199
## 9 A Wider Look at Visual Discomfort

9.1 Introduction ................................................. 202
9.2 Method ......................................................... 205
9.3 Results ....................................................... 210
9.4 Discussion ..................................................... 227

## 10 Migraine

10.1 Introduction ................................................. 229
10.2 Features of Migraine ......................................... 230
10.3 Migraine Aura .................................................. 231
10.4 Hyperexcitibility .............................................. 236
10.5 A Lack of Inhibition ......................................... 241
10.6 Elevated Internal Noise ....................................... 247
10.7 Habituation Deficits ......................................... 250
10.8 Limitations of the Explanations .............................. 253
10.9 Synchronisation of Neural Responses ........................ 255
10.10 Conclusion .................................................... 261

## 11 Conclusion

11.1 Visual Discomfort ............................................. 263
11.2 Neural Coding Efficiency ..................................... 263
11.3 Inadequate Accommodation Cues ............................ 265
11.4 Impact of Visual Discomfort ................................ 267
11.5 The Role of Contrast ........................................ 268
11.6 Conclusion ................................................... 268

## 12 Appendix

12 Appendix ......................................................... 270
LIST OF FIGURES

1.1 Amplitude spectra of noise and natural images .............. 7
2.1 Amplitude spectra ........................................... 21
2.2 Example stimuli .............................................. 22
2.3 Contrast matching results ................................... 24
2.4 More contrast matching ....................................... 25
2.5 2AFC individual discomfort judgements ...................... 29
2.6 2AFC individual discomfort judgements ...................... 30
2.7 2AFC discomfort judgements ................................ 30
2.8 2IFC contrast matching results .............................. 34
2.9 2IFC discomfort judgements ................................ 36
2.10 Printed stimuli discomfort judgements ...................... 39
3.1 Difference of Gaussian filter ................................. 46
3.2 Gabor filter ................................................... 46
3.3 Example Gabor filters ........................................ 47
3.4 ‘Bump’ stimuli ................................................. 50
3.5 Filtered noise stimuli of varying slope exponent, these are 1.5,
  2, 2.5 exponent values. ........................................ 50
3.6 Sine-wave striped stimuli. ................................... 51
3.7 Square-wave striped stimuli. ................................. 52
3.8 Riloid stimuli. ............................................... 53
3.9 ‘Blue Sun’ Image supplied by Akiyoshi Kitaoka at the College
  of Letters, Ritsumeikan University, Kyoto (Japan). ........ 54
3.10 ‘Clarch01’ Image supplied by Akiyoshi Kitaoka at the College
  of Letters, Ritsumeikan University, Kyoto (Japan). ........ 55
LIST OF FIGURES

3.11 ‘Enigma 780’ Image supplied by Michael Bach at the Augenklinik, University of Freiburg, Germany. Image originally created by Isia Leviant (1984). ........................................ 56
3.12 ‘Gira 3’ Image supplied by Akiyoshi Kitaoka at the College of Letters, Ritsumeikan University, Kyoto (Japan). .................. 57
3.13 ‘Warp’ Image supplied by Akiyoshi Kitaoka at the College of Letters, Ritsumeikan University, Kyoto (Japan). .................. 58
3.14 ‘Warp 7’ Image supplied by Akiyoshi Kitaoka at the College of Letters, Ritsumeikan University, Kyoto (Japan). .............. 59
3.15 Response kurtosis with increasing level of processing: original image kurtosis, DOG (LGN) and Gabor (cortex) filters. ...... 60
3.16 Response kurtosis ........................................ 62
3.17 Response kurtosis ........................................ 63
3.18 Mode response ....................................... 65
3.19 Mode response ....................................... 66
3.20 Mode response ....................................... 67
3.21 Mode response ....................................... 68
3.22 Cortical response distributions ................... 73
3.23 Cortical response distributions ................... 74
3.24 Cortical response distributions ................... 75
3.25 Cortical response distributions ................... 76
3.26 Cortical response distributions ................... 77
3.27 Cortical response distributions ................... 78
4.1 Amplitude spectra of natural images. .............. 81
4.2 Stimulus luminance profiles ....................... 88
4.3 Contrast matching results for Experiment 1 ......... 89
4.4 Discomfort scores for Experiment 1 ............... 90
4.5 Presented maximum contrast gradients .......... 92
4.6 Blur judgements for Experiment 2 ............... 94
4.7 Discomfort judgements for Experiment 3 ......... 98
4.8 Amplitude spectra of natural images. .......... 99
4.9 Modelled change in contrast gradients .......... 104
4.10 Maximum contrast gradient of phase manipulated stimuli .... 105
4.11 Natural image amplitude spectrum (log) change with defocus blur106
4.12 Sine wave radial waveform amplitude spectrum (log) change with defocus blur. ........................................ 107
4.13 Phase manipulation amplitude spectrum (log) and defocus. ... 108
4.14 Amplitude spectra. ................................... 109
<table>
<thead>
<tr>
<th>Section</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1</td>
<td>Shallow and large depth-of-field</td>
<td>116</td>
</tr>
<tr>
<td>5.2</td>
<td>Stereoscope setup</td>
<td>119</td>
</tr>
<tr>
<td>5.3</td>
<td>Object layout</td>
<td>121</td>
</tr>
<tr>
<td>5.4</td>
<td>Discomfort judgements against decreasing aperture</td>
<td>126</td>
</tr>
<tr>
<td>5.5</td>
<td>Discomfort judgements against decreasing aperture</td>
<td>129</td>
</tr>
<tr>
<td>5.6</td>
<td>Accuracy for all observers</td>
<td>130</td>
</tr>
<tr>
<td>6.1</td>
<td>Filtered noise search backgrounds</td>
<td>138</td>
</tr>
<tr>
<td>6.2</td>
<td>Average contrast thresholds</td>
<td>141</td>
</tr>
<tr>
<td>6.3</td>
<td>Results of Experiment 1</td>
<td>143</td>
</tr>
<tr>
<td>6.4</td>
<td>Striped search backgrounds</td>
<td>146</td>
</tr>
<tr>
<td>6.5</td>
<td>Results with small target</td>
<td>147</td>
</tr>
<tr>
<td>6.6</td>
<td>Results with large target</td>
<td>148</td>
</tr>
<tr>
<td>6.7</td>
<td>Results with horizontal target</td>
<td>151</td>
</tr>
<tr>
<td>7.1</td>
<td>The ‘wallpaper effect’</td>
<td>157</td>
</tr>
<tr>
<td>7.2</td>
<td>Serifs</td>
<td>160</td>
</tr>
<tr>
<td>7.3</td>
<td>Fonts</td>
<td>163</td>
</tr>
<tr>
<td>7.4</td>
<td>Morphed text</td>
<td>164</td>
</tr>
<tr>
<td>7.5</td>
<td>Actual text</td>
<td>164</td>
</tr>
<tr>
<td>7.6</td>
<td>The autocorrelation function of text</td>
<td>165</td>
</tr>
<tr>
<td>7.7</td>
<td>Acuity test for reading task</td>
<td>167</td>
</tr>
<tr>
<td>7.8</td>
<td>Number of mistakes when reading manipulated text</td>
<td>168</td>
</tr>
<tr>
<td>7.9</td>
<td>Average reading speed of manipulated text</td>
<td>169</td>
</tr>
<tr>
<td>7.10</td>
<td>Correlations between reading speed, discomfort and comprehension</td>
<td>171</td>
</tr>
<tr>
<td>7.11</td>
<td>Reading speed (wpm) against text stripiness for two fonts</td>
<td>175</td>
</tr>
<tr>
<td>8.1</td>
<td>Stimuli for Experiment 1</td>
<td>183</td>
</tr>
<tr>
<td>8.2</td>
<td>Comfort and clarity ratings against contrast</td>
<td>184</td>
</tr>
<tr>
<td>8.3</td>
<td>Text stimulus for the contrast experiments</td>
<td>187</td>
</tr>
<tr>
<td>8.4</td>
<td>Clarity and comfort ratings of striped patterns</td>
<td>189</td>
</tr>
<tr>
<td>8.5</td>
<td>Clarity and comfort ratings for natural images</td>
<td>190</td>
</tr>
<tr>
<td>8.6</td>
<td>Clarity and comfort ratings for text stimuli</td>
<td>191</td>
</tr>
<tr>
<td>8.7</td>
<td>Clarity and comfort judgements of striped stimuli</td>
<td>195</td>
</tr>
<tr>
<td>8.8</td>
<td>Clarity and comfort ratings of natural images</td>
<td>197</td>
</tr>
<tr>
<td>8.9</td>
<td>Clarity and comfort ratings of text</td>
<td>198</td>
</tr>
<tr>
<td>9.1</td>
<td>Bump stimuli for the online survey</td>
<td>207</td>
</tr>
<tr>
<td>9.2</td>
<td>Stripe stimuli</td>
<td>207</td>
</tr>
<tr>
<td>9.3</td>
<td>Harmonic-manipulated radial waveform stimuli</td>
<td>209</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

9.4 Optically blurred natural images. .......................... 209
9.5 Optically blurred filtered noise. .......................... 209
9.6 Phase manipulated radial waveform stimuli. ................. 210
9.7 Discomfort judgements. .................................. 212
9.8 Interest judgements. ................................. 213
9.9 Liking judgements. ........................................ 214
9.10 Magnification judgements. ............................. 215
9.11 Comparing discomfort across stimulus class ............... 217
9.12 Comparing interest across stimulus class .................. 218
9.13 Comparing liking across stimulus class .................... 219
9.14 Comparing magnification judgements across stimulus class 221
9.15 Discomfort judgements correlated with VDS, raw PS score and age. ............................... 223
9.16 Discomfort judgements correlated with VDS, raw PS score and age. ............................... 223
9.17 Interest judgements correlated against VDS, raw PS score and age. ............................... 224
9.18 Discomfort judgements correlated against VDS, normalised PS score and age. ............................... 225
9.19 Interest judgements correlated against VDS, normalised PS score and age. ............................... 226
10.1 A diagram of a cell ........................................ 238
10.2 Cell architecture ........................................ 241
CHAPTER 1

INTRODUCTION

1.1 Visual Discomfort

Visual discomfort is defined as the unpleasant sensations reported on viewing particular stimuli, such as high contrast striped patterns (Wilkins, Nimmo-Smith, Tait, McManus, Sala, Tilley, Arnold, Barrie, and Scott, 1984), abstract noise patterns (e.g. Fernandez and Wilkins, 2008; Juricevic, Land, Wilkins, and Webster, 2010), and text (e.g. Wilkins and Nimmo-Smith, 1987; Conlon, Lovegrove, Chekaluk, and Pattison, 1999; Nahar, Sheedy, Hayes, and Tai, 2007; Chase, Tosha, Borsting, and 3rd Ridder, 2009; Tosha, Borsting, 3rd Ridder, and Chase, 2009). Unpleasant sensations reported are wide ranging and show much individual variation: some individuals chosen from the general population report much discomfort, others report less, on viewing stimuli such as text (e.g. Conlon et al., 1999) and striped patterns (Evans and Stevenson, 2008). Symptoms have been reported to include headache, dry eyes, diplopia, and blur (Conlon et al., 1999; Sheedy, Hayes, and Engle, 2003), and illusory effects such as shadowy shapes, colours and scintillation (Evans and Stevenson, 2008).

Discomfort is necessarily subjective. Lambooij, Wijnand, Isselsteijn, and Hendrickx (2007) make the distinction between subjective feelings of discomfort and visual fatigue, which has objectively measurable consequences in terms of eye movements for example. However, in work by other authors the distinction between discomfort and fatigue can be somewhat blurred, with some reports using the terms interchangably (e.g. Shibata, Kim, Hoffman, and Banks, 2011). It seems likely that the two might share a common cause, and it is even possible in some cases that fatigue is the cause of discomfort. Results suggest that ocular motor difficulties might be a cause of discomfort in the case of visual display
1.1. **VISUAL DISCOMFORT**

units (VDS) (e.g. Nyman, Knave, and Voss, 1985). Additionally, both fatigue and discomfort might occur under stereoscopic presentation techniques. Under stereoscopic conditions, the object disparity specifies a distance from the observer that is different to that specified by accommodative (focussing) cues: the screen. This creates a cue conflict between depth cues provided by disparity and accommodation responses. As vergence and accommodation are usually yoked (Fincham, 1951) this is a possible cause of discomfort. A detrimental effect on the accommodative response has been demonstrated under these circumstances (MacKenzie, Hoffman, and Watt, 2010), as well as reports of subjective discomfort (e.g. Hoffman, Girshick, Akeley, and Banks, 2008; Shibata et al., 2011).

**Properties of Uncomfortable Stimuli**

The method of presentation is one possible way in which discomfort can occur through cue conflict. However, some stimuli are inherently uncomfortable, such as striped patterns (e.g. Wilkins et al., 1984). Discomfort from such patterns cannot be accounted for in terms of cue conflict. Discomfort must instead arise from their spatial properties, and how they are processed by the visual system. Therefore, the main focus of this study is to investigate visual discomfort in intrinsically uncomfortable stimuli, by manipulating their spatial properties, and to relate this to theories of visual processing.

There is a wide range of stimuli that can potentially evoke adverse effects in the observer. These stimuli include certain kinds of periodic patterns such as high contrast striped patterns (e.g Wilkins et al., 1984), filtered noise patterns (e.g. Fernandez and Wilkins, 2008; Juricevic et al., 2010), particular flickering patterns (e.g. Binnie, Findlay, and Wilkins, 1985), and even text (e.g. Nahar et al., 2007). Patterns evoking effects such as discomfort, headaches or motion illusions can be seen outside of the laboratory in escalator treads (Cohn and Lasley, 1990), Op-Art artworks (Zanker and Leonards, 2006) and paintings inspired by migraines, such as the work of Debbie Ayles (Fernandez and Wilkins, 2008). Additionally, individuals drawn from student populations have also reported discomfort on prolonged reading of text patterns (Nahar et al., 2007; Tosha et al., 2009; Chase et al., 2009). As well as periodic patterns, filtered noise patterns with differing spatial frequency content have been shown to be uncomfortable (e.g. Fernandez and Wilkins, 2008; Juricevic et al., 2010). These filtered noise patterns are not periodic stimuli, yet still evoke discomfort judgements based on their spatial attributes.
1.1. VISUAL DISCOMFORT

Certain spatial properties such as spatial frequency, contrast, size of pattern, etc affect the amount of discomfort reported (Wilkins et al., 1984). Contrast has differential effects depending on the type of stimulus: high contrast gratings are judged uncomfortable (Wilkins et al., 1984), but low contrast text is more uncomfortable than high (Nahar et al., 2007). Spatial frequency content is another property affecting discomfort judgements of both periodic stimuli and filtered noise patterns (e.g. Wilkins et al., 1984; Fernandez and Wilkins, 2008; Juricevic et al., 2010). Investigating the precise spatial attributes of uncomfortable patterns is useful in ascertaining why these stimuli cause discomfort. For example, it has been suggested that uncomfortable stimuli might cause excessive activity in the brain (e.g. Juricevic et al., 2010). Therefore it is logical that a high contrast stimulus might evoke even more of a response than a low contrast stimulus, thus exacerbating the problem. By using a theoretical framework, it is possible to characterise uncomfortable images and account for their effects in terms of visual processing.

Reasons for Studying Discomfort

There are many reasons for studying visual discomfort. One reason is to account for the effects in terms of visual processing, thus expanding our understanding of this system. In particular, there is an increasing body of research into optimal visual processing, whether cortical (e.g. Field, 1994), or ocular motor (e.g. MacKenzie et al., 2010). In both of these examples, evidence is provided to suggest that the visual system requires the stimulus to have certain properties in order for the visual system to process it effectively. A better understanding of the properties that make a stimulus suboptimal for the visual system might therefore give insights into how the system works.

Another reason for studying visual discomfort is that uncomfortable visual stimuli, particularly gratings, can have severe consequences for clinical populations typically associated with increased sensitivity to visual stimuli, such as migraine and epilepsy sufferers. For example, it has been reported that epileptogenic activity has been elicited in epilepsy sufferers on viewing striped patterns (e.g. Wilkins, Andermann, and Ives, 1975; Binnie et al., 1985). It has also been reported that in some cases seizures have been elicited by visual stimuli in epilepsy suffers (Radhakrishnan, Louis, Johnson, McClelland, Westmoreland, and Klass, 2005). Migraine sufferers report particular discomfort on viewing high contrast gratings (Marcus and Soso, 1989).

The consequences of viewing uncomfortable stimuli are not limited to clinical
1.1. VISUAL DISCOMFORT

populations, but can also impact those without a clinical diagnosis. There is also a spectrum of discomfort reported by those without a clinical diagnosis (e.g. Conlon, Lovegrove, Hine, Chekaluk, Piatek, and Hayes-Williams, 1998; Conlon et al., 1999; Conlon, Lovegrove, Barker, and Chekaluk, 2001). This is also evidence that this might affect performance on visual tasks (e.g. Conlon and Humphreys, 2001). This highlights the effects that visual stimuli can have on the brain, and the importance of studying these effects.

There are also reports of more serious consequences for non-clinical populations. For example, in the late nineties there were reports of the hospitalisation of over 600 children who watched a particular episode of a popular cartoon. This cartoon included large area of the screen that alternated blue and red at 12Hz (Ishida, Yamahita, Matsuishi, Ohshima, Ohshima, Kato, and Maeda, 1998). Temporally periodic stimuli are known to evoke epileptogenic activity (e.g. Binnie et al., 1985; Wilkins, Bonnanni, Porciatti, and Guerrini, 2004). However, wavelength of light is also a factor influencing discomfort in epilepsy sufferers (Main, Vlachonikolis, and Dowson, 1997). It has been shown that reddish light, of wavelength around 700nm being the most epileptogenic (Wilkins et al., 2004). Reddish light has also been chosen as most uncomfortable hue by migraine sufferers (Chronicle and Wilkins, 1991). Therefore it is unclear whether the temporal attributes, or the wavelength of the light was the cause of the disturbance.

Although this work opens up possibilities for attempting to prevent the consequences of uncomfortable stimuli, intervention methods are not directly addressed here. As stated above, another benefit of studying visual discomfort is a better theoretical understanding of visual processing. Developing a theoretical explanation to account for discomfort effects is therefore one of the aims of this study. One of the chief arguments of this thesis is that uncomfortable stimuli are those that are difficult for the visual system to process. Systematically investigating attributes of stimuli that the visual system has difficulty with could give insights into the way in which information is processed.

Multiple Causes of Discomfort

Stimuli could pose a difficulty for the visual system over any number of its processes. As seen above, discomfort could be due to difficulty from accommodation-vergence cue conflict. This would pose a potentially different cause to other potential sources of discomfort, such as discomfort from fluorescent lighting, which has also been reported (Conlon et al., 1999). It is thought that fluo-
1.1. VISUAL DISCOMFORT

Rescent lighting might cause discomfort either by glare from high luminance levels when combined with daylight, or due to the temporal frequency of the flicker associated, which might be subconsciously perceptible by some individuals (Winterbottom and Wilkins, 2009). There is evidence to suggest that the range of reported symptoms might be the results of distinct causes (Sheedy et al., 2003). Sheedy et al. (2003) investigated the reported effects of prolonged (15 minutes) viewing of text. The results were analysed using PCA and it was found that reported effects split into two main groups, which the authors denoted as those relating to external and those relating to internal causes. According to these authors, external effects were those relating to the surface of the eye, including dry eyes, which the authors suggested might be due to insufficient blinking. Internal effects included reports of headache, eyestrain, blurred vision, diplopia, and pain behind the eyes. Here the focus is on internal factors only: this work is concerned with how the visual system processes information, and not purely mechanical difficulties.

As there are so many potential aspects of the visual system that could have difficulty with stimuli, it is not possible to investigate them all in this thesis. Therefore the focus of this thesis is to investigate the contribution of two theoretical causes of discomfort: one relating to how visual stimuli are processed in the brain, and the other relating to the effectiveness of the visual system to use these stimuli to drive accommodation (focussing) responses. It is thought that particular images might not have the statistics expected by the visual system, and therefore not lend themselves to easy processing, either at a cortical level, or in terms of ocular motor responses. Thus the investigation is aimed at assessing how efficiently uncomfortable stimuli are processed in the brain, and the usefulness of the stimuli to drive accommodation.

These two explanations can be applied to some of the stimuli identified as uncomfortable. For example, individuals report discomfort from text (Conlon et al., 1999). It has been suggested that periodic stimuli might cause overly large cortical responses (Wilkins et al., 1984), and it has been suggested that text can be thought of as a periodic pattern (Wilkins, Smith, Willison, Beare, Boyd, Hardy, Mell, Peach, and Harper, 2007). Alternative theories involving ocular motor responses have been proposed as a cause of the discomfort from text patterns (Chase et al., 2009; Tosha et al., 2009). Therefore choice of stimulus is critical here to discern between the two theoretical explanations.

The kinds of stimuli that are optimal for the visual system might be those with the attributes typical of natural images. To understand why uncomfortable
stimuli might be difficult to process they must be compared to the processing of natural images. It is possible that uncomfortable stimuli present difficult stimuli for the visual system as they have spatial properties very different from those that typify natural scenes. The properties of natural images will be discussed in the next section.

1.2 Natural Images

Efficient visual processing of images depends on their statistical properties. Natural images are scenes typically seen in everyday environments, such as woods, beaches, vegetables, the office, etc. Some investigators distinguish between scenes containing man-made objects and those that do not (e.g. Torralba and Oliva, 2003). Torralba and Oliva (2003) showed that artificial environments typically contain more anisotropies based on the cardinal axes than images without man-made material. In other words, images of man-made environments have more oriented information based around the horizontal and vertical than completely natural images (e.g., woods etc), which are more isotropic (information at all orientations). Additionally, Torralba and Oliva (2003) showed that it is possible to classify images into those containing man-made objects and those without based on their amplitude spectra. However this distinction has not been upheld by other investigators (e.g. Baddeley and Hancock, 1991), whose work is relevant to the efficient coding models explained in more detail below. As the efficient coding models do not make this distinction between purely natural images, and those containing artificial material, this distinction will not be made in this thesis. Therefore the term ‘natural images’ includes all images that are of real-world scenes in this thesis.

Natural images, whether or not they contain man-made objects, are predictable - they are statistically very similar, occupying only a narrow range of all possible image statistics (Simoncelli and Olshausen, 2001). The visual system might exploit this predictability in order to process them with minimal effort (Simoncelli and Olshausen, 2001).

Black and white images can be considered as consisting of information contained as variation in contrast across space. Images can be analysed for their statistical properties using the Fourier Transform. This commonly used mathematical technique decomposes images into their spatial frequency content. Spatial frequency is a measure of contrast variation over space. In a simple grating, spatial frequency is the number of stripes in a given amount of space:
1.2. NATURAL IMAGES

Figure 1.1: Amplitude spectra of noise and natural images

high spatial frequency would thus consist of narrow stripes (rapid change in luminance over space), low spatial frequency of wider stripes. In natural images, this might be thought of as level of detail: finer detail is higher spatial frequency. In any image, there is a certain amount of contrast amplitude at each level of spatial frequency. This is the Fourier amplitude spectrum of the image. In white noise images, there is equal amplitude at all spatial frequencies, thus the amplitude spectrum is flat, see the left hand side of Figure 1.1.

In natural images, contrast amplitude typically falls with increasing spatial frequency. Thus one of the statistical properties typifying natural images is the relative amounts (contrast amplitude) of spatial frequency information present in the image: the $1/f^\beta$ amplitude spectrum, where $f$ is spatial frequency. The amplitude spectrum typical of natural images can be seen on the right hand side of 1.1. Contrast energy falls with approximately $1/f^3$, where $\beta$ is in the range 0.8 to 1.5 (Tolhurst, Tadmour, and Chao, 1992).

The amplitude spectrum is only one component of the Fourier Transform, the other component is the phase spectrum. The amplitude spectrum is a measure of how much information is present at each spatial frequency. The
phase spectrum is a measure of how the energy at different spatial frequency components lines up. Step edges contain many spatial frequency components. Components in phase are all aligned, and additive energy at many spatial frequencies in the same location creates the edge. Shifting the relative phases of the spatial frequency components will smudge the edge information over a larger area. Therefore the edge is no longer a step edge. The phase spectrum will be explained in more detail in the section on blur.

1.3 Sparse Coding

Natural image statistics have particular consequences for the visual system. Laughlin (1983) suggested that a prerequisite for efficient coding is that there must be good agreement between the characteristics of the code and the type of information it is representing. Central to this concept is the idea of redundancy. If there are relationships between the bits of information, then this is predictable. For example, English is not unpredictable. The sentence ‘The cat is ...’ is somewhat predictable. It is more likely that the missing word will be an adjective than a noun, for example. If this is the case, then the relationships between the words means the number of possible words to fill the gap is constrained. The same concept can be applied to images. For example, if it is more likely that a dark pixel will be next to a dark pixel than a white one, much of the image is actually predictable. Natural images tend to be predictable - contrast varies smoothly across surfaces, but not at edges. Edges are relatively rare, but contain much of the useful information of the image (Marr and Hildreth, 1980). If the image is predictable like this, then assumptions can be made about it to limit the number of units needed to code it - i.e. it might be possible to code only the (relatively infrequent) edges. For example, if the value of a certain number of pixels are known across the image then the unknown pixel values between them can be filled based on the assumption. In this manner, predictability of redundant images can be exploited by using a coding method that minimises the number of units needed to transmit the information in the image. However, if the image is totally random, then the same number of units are needed to code the image as are pixels present in the image itself. Thus if the image is not redundant, and therefore not predictable, no saving can be made on the number of units needed to process it. A sparse code can be used to reduce processing costs, but only if the input image matches the assumptions of the system.

In information theory, the total amount of information in a signal can be
divided into small chunks known as bits. If the information is predictable, then it is not necessary to transmit all possible bits for the message to be conveyed with sufficient accuracy. Instead, it is possible to transmit only some bits, and reconstruct the message on arrival. A measure of the predictability of information is the entropy. Predictable information has low entropy, meaning the information is largely redundant. Maximising the entropy of each transmitted bit of information is a scheme for conveying more information with fewer units, called compression. By increasing the entropy, more independent chunks of information are conveyed per unit, thus fewer units are required to convey the total message. Transmitting a compressed message, and recovering it on arrival is more efficient than using many units to code each small chunk of information separately.

In this manner, the visual system might be able to exploit the redundancy in the signal from images, if they are predictable. By maximising the signal entropy, the visual system can transmit sufficient information without needing so many active units, in this case neurons. However, this method depends on there being predictability in the information to be encoded. If the information is random, there is no redundancy to exploit, and thus no improvement in efficiency. If all bits of information are unrelated to each other then they must all be coded for individually. Many neurons would be needed to code the random stimulus; therefore it would have a high metabolic cost. Fortunately, natural images are not random; they have predictable characteristic properties, such as a $1/f$ amplitude spectrum (Tolhurst et al., 1992).

It is important that visual information is processed efficiently as it is metabolically expensive for neurons to be active. Some neurons in the visual cortex respond to visual stimuli by increasing their firing rate above a baseline level. This increase in activity demands more energy than when the neuron is at rest (the baseline response). Thus it is advantageous for neurons in the brain to process images efficiently as active units are metabolically expensive (Barlow, 1961). Lennie (2003) speculated that only 1/50th of all neurons can be active at any one time, and it would be impossible to supply enough energy for them to all to respond simultaneously. Thus there is a need for efficient allocation of metabolic resources. One possible way of efficiently coding stimuli might involve sparse coding.

Theoretical work has suggested that the visual system is optimised to code images with $1/f$ statistics efficiently, by using a sparse code (Field, 1987, 1994, 1999; Olshausen and Field, 2004). Sparse responses are desirable as only a few
units would be strongly active at any one time, and the rest would not respond. In this manner the visual system would conserve energy. A sparse code would be marked by a highly kurtotic response. Kurtosis is a measure of peakiness of a distribution, therefore a highly kurtotic distribution could mean that most neurons are inactive, and a sharp peak of a few neurons being strongly active. Highly kurtotic, sparse responses have been proposed as an efficient way of maximising signal transfer with minimal cost (e.g. Field, 1994, 1999). It is thought that the kurtosis of the response increases with progressive processing as the information is transferred from the retinal image to the lateral geniculate nucleus (LGN) and then to the cortex (Olshausen and Field, 2004).

Sparse coding has the advantage of high signal to noise ratios (Field, 1987, 1994; van Hateren and Ruderman, 1998; Field, 1999), as noise in the signal has been filtered out, which could result in good metabolic efficiency. Sparse codes have also been suggested to be an optimal response in terms of balancing the demands of information transfer and metabolic expense (Graham, 2006). Additionally, the theory of sparse coding is supported by cell recordings in the primary visual cortex (Vinje and Gallant, 2000). Critically, it has been suggested that sparse coding might be the optimal coding strategy for $1/f$ stimuli over multiple scales (Olshausen and Field, 2000), which are typical of the scale-invariant $1/f$ natural image statistics.

Redies, Hasenstein, and Denzler (2007) argued that a visual input that results in a sparse response would be aesthetically pleasing. Visual discomfort might be the reverse; an inefficient coding strategy to a particular stimulus would not result in a sparse response and therefore result in visual discomfort. Images with different statistics, for example those consisting of randomly generated noise will not become more sparse as they progress through the visual system, as they do not increase kurtosis with subsequent processing. Thus the relative amounts of spatial frequency information in the initial image is important to facilitate optimal coding.

It has been suggested that visual discomfort might be due to overly large cortical responses (Juricevic et al., 2010). This is possible as uncomfortable images are those with statistics different from those of natural images (e.g. Fernandez and Wilkins, 2008; Juricevic et al., 2010; O’Hare and Hibbard, 2011). As it is critical for a sparse code that the image statistics match the assumptions of the system, inefficient coding of these stimuli could well result in overly large responses. It has been shown that clinical populations show large cortical responses to periodic stimuli (e.g Wilkins, Darby, and Binnie, 1979). It has
been suggested that these large responses, to a less extreme extent, might be the cause of discomfort in non-clinical populations (Wilkins et al., 1984). There is also neuroimaging evidence that there is more activity on viewing uncomfortable stimuli than less uncomfortable stimuli (Huang, Zong, Wilkins, Jenkins, Bozoki, and Cao, 2011). Thus it seems a possibility that discomfort is due to inefficient coding of stimuli whose statistics do not match those the visual system is optimised to encode.

1.4 Accommodation Responses

Neural coding is not the sole facet of the visual system requiring specific input characteristics in order to respond optimally. Other aspects of the visual system also rely on appropriate stimulus properties for optimal performance. For example, the relative amplitude across spatial frequency is an important cue for accommodation (focussing) responses. It has been suggested that the accommodative system operates to maximise the contrast on the retinal image (MacKenzie et al., 2010). Indeed, when the contrast is too low the accommodative response remains at baseline levels, implying that the visual system makes no attempt to focus the image (Kotaluk and Schor, 1987).

The accommodative response is thought to be driven by a variety of cues, including feedback derived from the spatial frequency content of the image itself (Charman and Tucker, 1977). Low spatial frequency information is thought to drive the gross response, which is fine-tuned using the higher spatial frequency information (Charman, 1979). Evidence suggests that midrange spatial frequencies, around 3-5 cycles/degree, are important for the accommodative system (Owens, 1980; MacKenzie et al., 2010). Thus it is likely that natural stimuli contain sufficient relative amounts of spatial frequency information at various scales to provide a useful cue to the accommodative response.

By contrast, images containing a relative excess of low spatial frequency information are perceived as blurred (Murray and Bex, 2010). It is possible that an image lacking in appropriate amounts of the correct spatial frequency information will lead to uncertainty in the accommodation response (Owens, 1980). Uncertainty in the accommodative response might manifest itself in terms of increased microfluctuations (Day, Gray, Seidel, and Strang, 2009a). This uncertainty could be the source of discomfort. Thus, in addition to the possibility that inefficient neural coding could be the source of the discomfort, it could also result from uncertainty in the accommodative response.
Individual Differences in Discomfort

As previously stated, not all individuals report the same amounts of discomfort from the same stimuli (e.g. Conlon et al., 1999; Evans and Stevenson, 2008). Clinical populations especially report excessive discomfort from periodic stimuli, more so than healthy volunteers (Marcus and Soso, 1989). One possible explanation for this is in terms of an overresponsive cortex in the clinical populations (Wilkins et al., 1979). This might extend into a spectrum of discomfort affecting more of the general population. For example, it could be the case that the individual differences in the general population could be accounted for in terms of differing levels of cortical excitibility.

There are also potential reasons for individual differences in discomfort due to differences in optics. For example, the accommodative response typically starts to fail after the age of 40 (Ostrin and Glasser, 2004). It might be the case that the better the optics an individual has, the less tolerance they have for blur. This is logical: if the optics are too poor to notice the blur, then this blur will go undetected and not cause a problem. There is evidence to suggest that myopes have less sensitivity to blur (Wang, Cuiffreda, and Vasudevan, 2006). However, other studies contest this, finding no difference in accommodation response between myopes and emmetropes (e.g. Chen, 2002; Cufflin and Mallen, 2008). Therefore it is expected that individuals will not all report the same discomfort, and this may be predictable for some stimuli by their optics.

These two suggestions are only some of the possibilities for individual differences. Obviously these are not exhaustive or exclusive. However, by developing a theoretical account clear predictions are made regarding the origins of some of the individual differences. This will be investigated in Chapter 9.

1.5 Summary

To summarise, the main focus of this thesis is to investigate the image properties that affect visual discomfort in non-clinical populations. This will be investigated using two aspects of a theoretical framework. The chief argument is that stimuli that are uncomfortable are those that the visual system is unable to process optimally. This is either in terms of inefficient neural coding or by providing inadequate cues to accommodation.

The prevalence and severity of visual discomfort in the general population...
was also investigated. Additionally, the impact of uncomfortable stimuli on performance in visual tasks was also assessed. This was to ascertain whether reported discomfort was severe enough to prove detrimental to performance, and to provide a practical measure of the implications of discomfort from visual stimuli.

Note Concerning Structure

This is an introduction of the main argument of this thesis. As this thesis covers a broad range of topics (e.g. natural image statistics, spatial frequency, blur, depth-of-field, accommodation-vergence conflicts, binocular co-ordination, visual search, contrast, migraine), this introduction does not contain a complete literature review of relevant material for all the topics covered. Instead, each of the chapters covers the relevant literature in its own introduction. A short summary of each of the chapters is as follows:

Chapter 2: Spatial Frequency
The impact of spatial frequency on visual discomfort is assessed using filtered noise images. Psychophysical experiments were conducted, controlling for differences in perceived contrast and eccentricity.

Chapter 3: Modelling Sparse Coding
The sparseness of responses of the early visual system to uncomfortable stimuli were simulated using a sparse coding model. Sparseness was assessed by analysing response kurtosis, mode value, and distribution shape of the model output.

Chapter 4: Blur
Spatial frequency content is also important to blur perception and the accommodative response. In this chapter, discomfort judgements of blurred stimuli were made by participants in psychophysical experiments. The change in contrast due to optics was modelled using a simplified model of the modulation transfer function of the human eye.

Chapter 5: Depth-of-field
Blur location is critical to its influence on the accommodative response. Whether the location of blur influenced discomfort judgements was assessed using depth-
of-field, which is peripheral and does not directly influence accommodation. The effects of conflict between depth-of-field blur and other depth cues (vergence and accommodation) was assessed.

Chapter 6: Visual Search
Observers have previously made discomfort judgements of certain stimuli. Whether stimuli previously established to affect discomfort judgements also influenced performance on a simple visual search task was investigated in this chapter.

Chapter 7: Reading
Text is a periodic stimulus, which could be uncomfortable due to binocular co-ordination difficulties or large cortical responses. Additionally, there is an established association between visual discomfort and reading performance. This study consisted of two experiments investigating the impact of text manipulations designed to facilitate binocular co-ordination or reduce the potential for large cortical responses.

Chapter 8: Contrast
The effects of contrast and luminance on discomfort judgements were assessed in this study. Participants rated images of striped patterns, text and natural images with contrast and luminance systematically manipulated.

Chapter 9: A Wider Look at Visual Discomfort
The prevalence of visual discomfort sufferers in a wider population than in the previous experimental chapters was investigated using an online survey. Respondents completed standardised measures of visual discomfort and rated images similar to those used in the previous experiments involved in this thesis.

Chapter 10: Migraine
Migraine is a disorder with associations with vision through visual symptoms and differences in visual performance. Additionally, greater visual discomfort is typically reported by migraineurs than controls. This chapter reviews the relevant literature to on migraineur neural responsiveness, performance on visual tasks, and the theoretical frameworks underlying the disorder.
CHAPTER 2

SPATIAL FREQUENCY AND VISUAL DISCOMFORT

2.1 Introduction

Many factors might influence perception of discomfort from visual stimuli. In this chapter the focus is on the role of spatial frequency as a determinant of discomfort. Using striped patterns of differing spatial frequency, Wilkins et al. (1984) assessed discomfort by asking participants about the number of illusions and aversive sensations they experienced on viewing these patterns. It was found that striped patterns of spatial frequencies within the range of 2-8 cycles/degree were particularly associated with visual discomfort, more than higher or lower spatial frequencies. This was investigated further by Fernandez and Wilkins (2008) who provided additional evidence for the importance of a similar range of spatial frequencies, this time those within two octaves of 3 cycles/degree. They compared the Fourier amplitude spectra of artworks judged as uncomfortable to view, with those for artworks that were judged as comfortable. They found a pronounced concentration of energy within two octaves of 3 cycles/degree in the former. This suggests that the presence of relatively large amounts of energy around this spatial frequency contributes to the experience of discomfort. This was tested experimentally by manipulating the amount of energy in this region in random noise patterns. It was found that increasing the amount of energy within two octaves of 3 cycles/degree increased ratings of discomfort, while reducing the amount of energy around this frequency decreased these ratings. As noted by Fernandez and Wilkins (2008), spatial frequencies around this range have been associated with the
2.1. INTRODUCTION

highest levels of aversion, and have a greater probability of inducing seizures in some epilepsy sufferers than other frequencies. For example, spatial frequencies between 1 and 4 cycles/degree were found to be optimum for causing paroxysmal activity in EEG recordings by Binnie and Darby (1980).

One suggestion as to why some images might be more uncomfortable than others is that discomfort is associated with deviations from the spatial statistics of natural images (Fernandez and Wilkins, 2008; Juricevic et al., 2010). Fernandez and Wilkins (2008) examined the effects of deviations from a $1/f$ amplitude spectrum. Images with an excess of contrast energy at frequencies around 3 cycles/degree were found to be more uncomfortable than images with a $1/f$ natural slope. These results suggest that deviations in the amplitude spectrum of images away from the $1/f$ natural slope might cause discomfort. Consistent with this view, Juricevic et al. (2010) showed that varying the slope of the amplitude function of images affected their rated discomfort. They used random noise images, and ‘Mondrian’ patterns comprising randomly positioned overlapping rectangles. They found the lowest discomfort ratings for a natural slope, with judged discomfort increasing for higher and lower slopes.

It has also been suggested that the aesthetic value of images is influenced by the extent to which they conform to the expected statistics of natural images (Redies et al., 2007; Graham and Redies, 2010). As it might be expected that aesthetics has a component associated with the pleasure of viewing a stimulus, it might be expected that images deviating from these $1/f$ statistics are judged as uncomfortable. From this overall body of evidence, a case is made that those images that do not conform to natural image statistics will cause discomfort.

The observation that uncomfortable images might be those that deviate from natural image statistics does not explain why such departures might be associated with discomfort. One possible explanation involves the idea of optimal processing of incoming visual stimuli. It is often assumed that the visual system is in some way optimised for images with properties that are typical of those encountered in the natural environment. It follows from this that there will be other types of images for which its responses will not be optimal. Discomfort might arise from this non-optimal response. One issue that has received considerable attention is whether the visual system is optimised for encoding images with a $1/f$ amplitude spectrum. The visual system is organised into spatial frequency channels. A channel is a collection of neurons that respond to a similar attribute - in this case, spatial frequency. Each channel is responsible for processing visual information in a certain spatial frequency range, called a bandwidth (see DeVlois, Albrecht, and Thorell, 1982, p. 177).
The bandwidth of each channel increases in octaves. Field (1994) argued that the bandwidths of spatial frequency tuned channels are optimised for this type of amplitude spectrum. Visual information is processed by channels that are tuned for spatial frequency. These channels can be characterised by their preferred spatial frequency, and their spatial frequency bandwidth. The latter refers to the range of frequencies to which they are sensitive. Measurements of the bandwidths of spatial frequency tuned channels have shown that, as the preferred frequency of a channel increases, its bandwidth increases when measured on a linear scale, but remains constant on an octave scale (Field, 1994). Stimuli with a $1/f$ amplitude spectrum have greater energy at low spatial frequencies than at high. The constant octave bandwidth of channels means that the total amount of information carried by each channel will be constant for images with a $1/f$ amplitude spectrum. Conversely, when stimuli do not have a $1/f$ amplitude spectrum, the amount of information carried will vary across channels. This imbalance across channels might cause discomfort by not being optimally processed, creating an overly large response in a particular population of neurons. The response would not be distributed sparsely across all possible processing resources, which would mean a less efficient response, thus discomfort might occur.

Another way in which the visual system might be optimised for $1/f$ statistics is in its differential sensitivity for different spatial frequency information, known as the contrast sensitivity function. It is well established that sensitivity to low-contrast, narrowband stimuli (such as gratings) peaks for frequencies around 4 cycles/degree, and decreases for both higher and lower frequencies (Campbell and Robson, 1968). The fall-off in sensitivity for low spatial frequencies can be seen as an optimal adaptation to $1/f$ images because, as contrast increases with decreasing frequency, signals may be accurately transmitted with a low gain, in order to minimise the metabolic cost with little loss of information. As signals have a higher amplitude at low frequencies in the first instance, it is beneficial to reduce the gain so that metabolic energy is not expended unnecessarily. At high spatial frequencies, there is little useful information as the signal amplitude is expected to be low. Gain here is therefore low in order to avoid expending energy on signals with a low signal-to-noise ratio (Atick, 1992), which would serve simply to amplify the noise. The gain in both cases is thus optimised to equalise the responses resulting from the particular statistics of the typical natural visual environment (i.e. the $1/f$ amplitude spectrum). Again, it follows that signals which do not have the expected $1/f$ amplitude spectrum will be encoded inefficiently. The consequences of this might include an unequal distribution of activity across spatial frequency channels, an inef-
2.1. INTRODUCTION

It has been proposed that discomfort arises when visual stimuli produce exceptionally strong neural responses (Wilkins et al., 1984; Wilkins, 1995). Concentration of energy around 3 cycles/degree might be expected to increase the strength of the neural response, since this frequency is around the peak of the contrast sensitivity function. The strength of the neural response will be expected to increase as the contrast of an image is increased. Stimulus contrast has been shown to affect discomfort judgements: higher contrast can lead to greater discomfort judgements in striped and filtered noise stimuli (Wilkins et al., 1984; Fernandez and Wilkins, 2008; Juricevic et al., 2010).

When assessing the effects of manipulating the Fourier amplitude spectrum on discomfort, Fernandez and Wilkins (2008) were careful to control for the physical (RMS) contrast of their stimuli. They did this in two different ways. In one set of stimuli, the amplitudes of the natural slope and bump images were matched at the 3 cycles/degree peak of the bump. This meant that the amplitude was lower in the bump images than in the natural slope images at all other frequencies (RMS contrast was 0.12 and 0.2 respectively), and they therefore had a lower amplitude overall. Another set of stimuli was created in which the overall contrast of the images was matched (RMS contrast of 0.2). They found the same pattern of discomfort judgements in both cases.

However, it is unclear whether observers would have perceived either set of stimuli as having the same contrast. Observers are more sensitive to spatial frequencies around 4 cycles/degree than they are to higher or lower frequencies (Campbell and Robson, 1968). Therefore, adding a bump in the amplitude spectrum around this spatial frequency might create a greater response compared to other stimuli with bumps at higher or lower spatial frequencies, even when stimuli are matched for physical contrast. This could then account for visual discomfort in terms of a disproportionately large response to excess contrast around the frequency of 3 cycles/degree suggested by Fernandez and Wilkins (2008). That is, the influence of spatial frequency on visual discomfort could be a contrast effect resulting from the differential sensitivity to different spatial frequencies, rather than a direct result of spatial frequency per se.

The contrast sensitivity function measures the detectability of stimuli at low (just perceptible) levels of contrast. There is evidence of a flattening of this function at suprathreshold contrast levels (Georgeson and Sullivan, 1975; Georgeson, 1985; Georgeson) and for stimuli more complex than the simple gratings...
used to determine the contrast sensitivity function (Bex and Langley, 2007). This phenomenon is known as contrast constancy. For example, two stimuli with a different spatial frequency but the same contrast appear more similar in contrast than would be predicted by a simple comparison of their contrast thresholds. This flattening is not, however, complete (Bex and Makous, 2002): stimuli for which contrast thresholds are lowest tend to appear higher in contrast at a given level of suprathreshold contrast. Therefore, when contrast amplitude in broadband stimuli is concentrated around spatial frequencies to which humans are particularly sensitive, this might result in an increase in their perceived contrast.

Differences in perceived contrast could potentially account for some of the effects reported by Fernandez and Wilkins (2008). When stimuli were matched for overall physical contrast, bump stimuli might be expected to have a higher perceived contrast then natural slope stimuli. Discomfort might then reflect these differences in perceived contrast. This explanation would be consistent with the view that discomfort can arise from cortical hyperexcitation; the increased response from the stimuli with higher perceived contrast might have been judged as more uncomfortable. It is not possible to predict exactly how perceived contrast will be affected by changes in the amplitude spectra for complex stimuli. Therefore, the aim of the first experiment was to measure perceived contrast in the current stimuli. This will allow an assessment of how perceived contrast varies with the location of the peak in spectral power. It will also allow the creation of stimuli that are matched for their perceived contrast, which may then be used in the experiments on visual discomfort.

A further complication is that the relationship between image contrast and discomfort is in fact non-monotonic (Juricevic et al., 2010; Fernandez and Wilkins, 2008). Discomfort increases when contrast is either too low (Nahar et al., 2007; Juricevic et al., 2010) or too high (Wilkins et al., 1984; Fernandez and Wilkins, 2008). These results are likely to reflect a combination of reduced visibility and a poor signal for accommodation at low contrasts, and strong neural responses at high contrasts (e.g. Wilkins, 2003). For example, Owens (1980) showed that sine-wave stimuli at the extremes of the contrast sensitivity function are poorer stimuli for accommodation. These effects might also depend on the type of stimulus: text stimuli are judged more uncomfortable at low contrast (Nahar et al., 2007), whereas discomfort reported from striped patterns increases with contrast (Wilkins et al., 1984). The strength of any effects might also be determined by individual differences in sensitivity (Conlon et al., 2001).
This first set of experiments was conducted to clarify the role of spatial frequency in visual discomfort in broad-band images. One particular aim was to dissociate such effects from potential differences in perceived contrast. The first aim was therefore to determine if broad-band noise stimuli that are matched for their physical contrast differ in their perceived contrast when contrast amplitude is concentrated at different spatial frequencies. It is possible that any such differences could account for the effects of spatial frequency on visual discomfort that have been found in previous studies (Wilkins et al., 1984; Fernandez and Wilkins, 2008). This, in itself, would represent an important insight regarding the factors contributing to discomfort. The second aim was to determine whether such effects are tuned for spatial frequency (i.e. whether particular spatial frequencies are judged more uncomfortable than others) or if all spatial frequencies are equally uncomfortable once such potential contrast effects are taken into account.

2.2 General Methods

Apparatus

Stimuli were presented at a distance of 1m, on a 21-inch Sony Trinitron monitor with a screen resolution of 1680 x 1050 and a vertical refresh rate of 60 Hz. A chin-rest was used to maintain the correct viewing distance. One pixel subtended 0.8 arcmin. Images were created and displayed using MATLAB (The Mathworks, Inc, 2005) and the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997). The luminance response of the monitor was measured and calibrated using a Minolta LS-110 photometer. The luminance of the mid-grey background was 38.5 cd/m². The luminance range was from 3 to 74 cd/m².

Stimuli

Stimuli were based on those in Study 3 of Fernandez and Wilkins (2008). 840 x 840 pixel Gaussian white noise images were filtered in the Fourier domain to produce images with a $1/f$ natural slope amplitude spectrum. The left hand side of Figure 2.1 plots the $1/f$ amplitude spectrum on log axes. The terminology of Fernandez and Wilkins was adopted, and therefore these stimuli are referred to as ‘natural slope’ images. For some of the stimuli, a peak was added to the amplitude spectrum by multiplying the $1/f$ spectrum with a raised radial cosine filter that was symmetrical in log axes:
2.2. GENERAL METHODS

\[ H(f) = \begin{cases} T \\ \frac{T}{T}[1 + \cos(\frac{\pi T}{T})(|\log(f) - \log(f_0)| - \frac{1-\beta}{2T})] \\ O \end{cases} \]

(2.1)

for \( (0 \leq |\log(f) - \log(f_0)| \leq \frac{1-\beta}{2T}) \)

\( (\frac{1-\beta}{2T} \leq |\log(f) - \log(f_0)| \leq \frac{1-\beta}{2T}) \)

\( (|\log(f) - \log(f_0)| > \frac{1-\beta}{2T}) \)

Where \( T \) is 0.9, \( \beta \) is the roll-off factor of 0.5, \( f \) is the spatial frequency, and \( f_0 \) is the centre frequency of the peak. These are referred to as ‘bump’ images. An example of a bump stimulus amplitude spectrum can be seen on the right hand side of Figure 2.1. Because the spatial frequency tuning of the effects of manipulating the amplitude spectrum in this way was of interest, a smaller bandwidth (0.5 octaves) than that used by Fernandez and Wilkins (2008) (3 octaves) was used. The relative amplitude at the peak spatial frequency was increased by a factor of 10. The peak spatial frequencies used were 0.188, 0.375, 0.75, 1.5, 3, 6, 12 and 24 cycles/degree. Examples of a natural slope stimulus and some of the bump stimuli are shown in Figure 2.2.

Figure 2.1: Amplitude spectra

Left: log spatial frequency (x-axis) plotted against log amplitude (y-axis) for a 1/f natural slope. Right: the amplitude spectrum of an example bump stimulus.
2.3. **EXPERIMENT 1: CONTRAST MATCHING**

All stimuli were multiplied by a window with a central circle subtending a diameter of 5.4° and a Gaussian drop-off at the edges (standard deviation of 0.93°). Each visible pattern subtended approximately 8.53°. On each trial, two stimuli were presented side-by-side at the same time. One was presented on the left, with its centre at an eccentricity of 5.71° relative to the centre of the screen. The other was presented at the same eccentricity on the right of the screen. The physical contrast of the stimuli was defined in terms of their RMS contrast, which has been shown to be the most representative measure of contrast for complex images (Peli, 1990);

\[
\sqrt{\frac{\sum (I - \bar{I})^2}{\sum I}}
\]  

(2.2)

Where \( I \) is the intensity of a pixel, normalised between 0 and 1, and \( \bar{I} \) is the average intensity. Ten versions of each stimulus were created, in order to prevent observers recognising particular patterns.

![Figure 2.2: Example stimuli](image)

Some of the stimuli used in the experiment (left to right): 0.75, 1.5 and 3 cycles/degree bump stimuli and natural slope stimulus.

### 2.3 Experiment 1: Contrast Matching

The aim of the first experiment was to measure the degree of contrast constancy for the bump stimuli used by Fernandez and Wilkins (2008). The first aim was to establish firstly whether bump stimuli have the same perceived contrast as natural slope stimuli that are matched for physical contrast, and secondly whether perceived contrast is affected by the peak spatial frequency of the bump; i.e. if all bump stimuli were judged to have the same discomfort. These
2.3. EXPERIMENT 1: CONTRAST MATCHING

results will be used to create stimuli that are matched for their perceived contrast, to be used in the discomfort rating experiments.

Method

Observers: Two experienced psychophysical observers took part in the experiment. Both had corrected-to-normal vision. The results of these observers were used to create the contrast matched stimuli. To ensure generalisability to the wider population, three naïve observers took part in a subsequent experiment. All had corrected-to-normal vision. All experiments were approved by the University of St Andrews University Teaching and Research Ethics committee.

Procedure: The task was a spatial two-alternative forced-choice (2AFC) procedure, using the method of constant stimuli. On each trial a standard and test stimulus were presented simultaneously, side-by-side on a mid-grey background. Presentation time was 1.8s, with an abrupt onset and offset (i.e. a rectangular temporal window). Between stimulus presentations, the screen was black. A central white fixation cross was present throughout. Observers pressed the left or right arrow key on the computer keyboard to indicate which stimulus appeared to have the higher contrast. The next trial was presented when the observer pressed a response key.

The standard was a natural slope stimulus with the same mean luminance varying around the value of the mid-grey background (luminance = 38.5 cd/m²). There were 10 versions of the \(1/f\) slope standard stimulus, with the mean RMS contrast of 0.059, and standard deviation of 0.001.

Each test set consisted of a bump at a different spatial frequency. Each spatial frequency in turn was manipulated to have a range of ten different contrasts, decided on by a pilot study. For the 0.375, 0.75, 1.5 and 3 cycles/degree spatial frequencies the range of contrast was from 0.6 to 1.5 times the RMS contrast of the standard, increasing in increments of 0.1. For 0.188, 6 and 12 cycles/degree stimuli the ranges were 1.1 to 2, 1.1 to 2 and 1.6 to 2.5 times the contrast of the standard, respectively. All stimuli had the same mean luminance. Each block compared only one spatial frequency to the natural slope standard. Each block consisted of 40 repetitions of each of the 10 different contrast stimulus comparisons, 20 with the standard on the left, 20 with the standard on the right. Each observer completed one such block for each of the eight spatial frequencies.
2.3. EXPERIMENT 1: CONTRAST MATCHING

Results

Observers’ responses were plotted as the proportion of trials in which the test was chosen as having more contrast than the standard against the contrast of the test stimuli. A cumulative Gaussian was then fitted to the data using the ‘psignifit’ toolbox (Wichmann and Hill, 2001a,b). The Point of Subjective Equality (PSE), defined as the level at which the observer chose the standard as the higher contrast on half the trials, was determined from the fitted curve. PSEs for the two experienced observers are plotted against spatial frequency in Figure 2.3. The highest spatial frequency tested (24 cycles/degree) is not included on this graph as it could not be matched for contrast; even at maximum achievable contrast the natural slope standard still appeared to have more contrast than this stimulus. The results of this experiment were used to match the stimuli for perceived contrast in Experiment 2.

Figure 2.3: Contrast matching results

Relative RMS contrast is plotted (y-axis) against spatial frequency of stimuli (x-axis). The U-shaped tuning function for PSEs for two observers, LOH and PBH, indicates that perceived contrast is not the same across spatial frequencies. Error bars show 95% confidence limits.

The individual results for the three naïve observers who also took part in this
2.3. EXPERIMENT 1: CONTRAST MATCHING

The results of the experiment are plotted in Figure 2.4. This figure also shows that the average of the two authors’ data (i.e. the values that were used to create stimuli matched for perceived contrast in the second experiment), and the average of all five observers. Despite some individual variation there is a clear U-shaped tuning function for all observers. Stimuli with the same RMS contrast, but different amplitude spectra, did not appear to have the same contrast. There also seems to be good agreement between the average data for the two authors and the average data across all observers.

![Figure 2.4: More contrast matching](image)

Relative RMS contrast is plotted (y-axis) against spatial frequency (x-axis) for the three extra observers. Individual variation can be seen in the U-shaped tuning functions across observers. Error bars show 95% confidence limits. The fourth plot is the average of all five observers (open circles) and the average of the original two experienced observers (filled circles). There is little difference in the shape of the tuning function for the averages. Again, perceived contrast is not the same across spatial frequencies.
Discussion

These results show that broadband bump stimuli of the kind used by Fernandez and Wilkins (2008) tend not to have the same perceived contrast as natural slope stimuli that are matched for RMS contrast. Also, there are variations in perceived contrast as a function of the central spatial frequency of the bump. There is individual variation in the exact spatial frequencies judged as having more contrast, for example observer two shows low pass tuning (see Figure 2.4). However, there is a broadly similar U-shaped tuning function for all observers (except for observer two), indicating that stimuli that are matched for physical contrast are not matched for perceived contrast. Furthermore, the middle of the range of spatial frequencies used seems to be generally perceived to have a higher contrast than the higher and (with the exception of observer two) lower spatial frequency images.

This raises the possibility that the findings of Fernandez and Wilkins (2008) might in part reflect differences in perceived contrast, which might in turn reflect greater cortical excitation for some stimuli. That is, concentration of energy around three cycles/degree might increase the perceived contrast of stimuli. The results clearly show that the stimuli do not have the same perceived contrast when matched for RMS contrast. These results will be used to compensate for the effects of perceived contrast in the second experiment, which investigates discomfort judgements.

2.4 Experiment 2: Relative Discomfort Judgements

The first aim of the second experiment was to determine whether bump stimuli are judged more uncomfortable than natural slope stimuli when they are approximately matched for their perceived contrast. The second aim was to determine whether discomfort ratings for these stimuli are affected by the spatial frequency of the peak of the bump. Perceived contrast was not matched for each participant individually, for a number of reasons. Firstly, this would have required the presentation of stimuli with a different physical contrast for each participant, which was to be avoided. Secondly, observers would have been exposed to many examples of the experimental stimuli before they were asked to judge their relative discomfort, which may have affected their responses. Finally, this would have significantly increased the length of the experiments.
2.4. EXPERIMENT 2: RELATIVE DISCOMFORT JUDGEMENTS

Stimuli were instead approximately matched by the average contrast matching function for the two authors. This manipulation had the effect of increasing the physical contrast of those stimuli that had a low perceived contrast, and decreasing the physical contrast of those that had a high perceived contrast.

Method

Observers: Eleven naïve observers, four male, seven female, with normal or corrected-to-normal vision, participated in the first condition in this experiment, and 15 in the second. All participants were students at the University of St Andrews, and all were over 18 years of age. All 11 participants from the first condition took part in the second condition. All participants gave their informed consent to participation.

Stimuli: Stimuli were created as before, but were matched for perceived contrast according to the results of experiment one by taking the mean of the PSEs for the two authors, and using stimuli with this RMS contrast. All of these stimuli should therefore have approximately the same perceived contrast as each other, and the natural slope standard, given the good agreement between observers in experiment one. Stimuli with a bump at 0.75, 1.5 or 3 cycles/degree therefore had a lower RMS contrast than the standard. The remaining stimuli (0.188, 0.375, 6 and 12 cycles/degree) had a higher RMS contrast. The natural slope stimulus remained with an RMS contrast of 0.059 and a mean luminance of 38.1 cd/m².

Procedure: There were two conditions. In the first condition the standard was the natural slope stimulus. In the second condition, the standard was the 3 cycles/degree bump stimulus. In both conditions, bump stimuli at all seven spatial frequencies tested were compared with the standard. This means that for some trials, the 3 cycles/degree stimulus was both the standard and the test stimulus. The second condition was conducted to provide a more direct test of any possible spatial-frequency tuning. It is possible, for example, that all bump stimuli would be judged as less comfortable than the natural slope standard, but this would not necessarily indicate their discomfort relative to one another. Within each block of trials, all seven bump stimuli were compared with the standard 30 times. This resulted in 210 trials per block. Stimuli were again presented for 1.8s, with an abrupt onset and offset. Two blocks of trials were run, one for each of the two standards. Each block took around 20 minutes to complete. As before, the test stimulus was presented on the right in half of the trials, and on the left in the other half. Ten different,
randomly chosen versions of each stimulus were used to avoid recognition of individual images. Participants chose the stimulus they considered to be more comfortable. This particular phrasing was used as it was felt more intuitive than asking observers to pick the uncomfortable image, since none of the images was particularly uncomfortable. For example, no high-contrast striped patterns were used. Observers responded by pressing the left or right arrow key on the computer keyboard.

Results

Each trial consisted of a presentation of two images, the test and the standard, and the observer was asked to judge which of the two was the more comfortable. The percentage of times a test stimulus was considered more uncomfortable than the standard was taken as a measure of discomfort. Figure 2.7 plots relative discomfort judgements (percentage of times a test stimulus is considered the less comfortable of the pair) as a function of the spatial frequency of the peak of the bump. As all observers showed a similar pattern of results, the plotted line is the average of all observers’ discomfort judgements. A sample of individual data from four observers is shown in 2.5 compared to a natural slope standard, and 2.6 compared to a bump standard.

The left hand side of Figure 2.7 plots results averaged across observers for the first condition, in which stimuli were compared with a natural slope standard. The right hand side shows the results for the second condition, in which stimuli were compared with a 3 cycles/degree bump standard.

When the standard was a natural slope stimulus, it tended to be judged as the more comfortable image. Eleven observers judged seven spatial frequencies of bump against natural slope stimuli, and in 70 out of these 77 cases the bump stimuli tended to be judged as the more uncomfortable. The probability of at least this number of judgements that the bump is the less comfortable stimulus by chance is less than $10^{-5}$. In general, bump stimuli were judged as more uncomfortable than natural slope stimuli that were matched for perceived contrast. The results also showed clear spatial frequency tuning; not all bump stimuli were equally likely to be judged as the more uncomfortable. As there were non-normal distributions of the data due to ceiling effects for some of the
2.4. EXPERIMENT 2: RELATIVE DISCOMFORT JUDGEMENTS

Figure 2.5: 2AFC individual discomfort judgements

Percentage of times that the test stimulus was chosen as more uncomfortable (y-axis) against spatial frequency (x-axis) for bump stimuli compared to the natural slope standard for four observers.

stimuli, non-parametric statistics were used. A Friedman’s test showed this effect of spatial frequency to be significant ($\chi^2(6) = 25.77, N = 11, p < 0.001$).

This spatial frequency tuning was tested more directly in the second condition, in which bump stimuli with different spatial frequencies were compared. On each trial, a test stimulus was compared with a 3 cycles/degree bump standard. The results of this condition are plotted on the right hand side of Figure 2.7. When the test was also a 3 cycles/degree bump (i.e. the same as the standard) preferences were at chance levels (50%), as would be expected. For other frequencies, the results show clear spatial frequency tuning. Again, results of a Friedman’s test showed this spatial frequency tuning to be significant: ($\chi^2(6) = 49.10, N = 15, p < 0.001$). This spatial frequency tuning was explored further by performing a Wilcoxon test, comparing the responses averaged across low spatial frequencies (0.188-1.5 cycles/degree) with responses in which the 3 cycles standard was compared with itself. This was also done for the average of the responses across the two highest spatial frequencies (6...
2.4. EXPERIMENT 2: RELATIVE DISCOMFORT JUDGEMENTS

Figure 2.6: 2AFC individual discomfort judgements

Percentage of times that the test stimulus was chosen as more uncomfortable (y-axis) against spatial frequency (x-axis) for bump stimuli compared to the bump standard for four observers.

Figure 2.7: 2AFC discomfort judgements

Left: Percentage of times that the test stimulus was chosen as more uncomfortable (y-axis) against spatial frequency (x-axis) for bump stimuli compared to the natural slope standard. The bump stimuli are more often considered more uncomfortable than the natural slope standard. Right: Results when bump stimuli were compared to a 3 cycles/degree bump standard. The lower spatial frequencies (1.5, 0.75, 0.375 cycles/degree) were more often judged more uncomfortable than the 3 cycles/degree standard. Error bars show ± 1 standard error.
and 12 cycles/degree). Bonferroni corrections for multiple comparisons were performed. Thus whether moving the peak of the energy to lower or higher spatial frequencies tended to increase or decrease judgements of discomfort could be determined. The median number of discomfort responses to low spatial frequency bumps (0.19 to 1.5 cycles/degree; median = 13) was significantly higher than those to the 3 cycles/degree test stimulus (median = 0) ($T(14) = 6, p < 0.05$). The median number of discomfort responses to stimuli with a bump at higher spatial frequencies (6 and 12 cycles per degree; median = -13) was not significantly different from responses to the 3 cycles/degree test stimulus (median = 0) ($T(14) = 31, NS$).

**Discussion**

The results of the first condition show that the filtered noise pattern with a $1/f$ natural slope amplitude spectrum is consistently considered more comfortable than a stimulus with a relative increase in energy at some point (i.e. a peak) in the amplitude spectrum (the bump stimuli). This finding cannot be explained by differences in perceived contrast, as stimuli were approximately matched for this. It is likely that stimuli would not be exactly matched for contrast for each individual, since the same contrasts were used for all observers. These were derived from the contrast matching data of the two experienced psychophysical observers. However, given the contrast matching data for three naïve observers in experiment one, these results appear to be representative of perceived contrast for the wider population. It is also instructive to note that discomfort judgements were not determined by physical contrast. Contrast was raised for both the lowest and highest spatial frequencies used, whereas discomfort judgements were not.

One interpretation of this finding is that deviation from the typical amplitude spectrum of natural images increases visual discomfort. However, this explanation alone cannot account for the clear tuning of the results for spatial frequency. All the bump stimuli deviated by the same amount from the statistics of natural images, and therefore there should be no difference between them if deviation from natural slope spectra were the sole cause of visual discomfort.

This spatial frequency tuning was assessed more directly in the second condition. Stimuli with a peak at a frequency lower than the three cycles/degree standard were consistently chosen as less comfortable. This is not exactly the same as the frequency associated with maximum discomfort by Fernandez and Wilkins (2008). They suggested that frequencies within 2 octaves of 3
cycles/degree should be the most uncomfortable. Whilst those lower than 3 cycles/degree are more uncomfortable than the three cycles/degree standard, those above (6 and 12 cycles/degree) were not. There are a variety of possible reasons for this. One possibility is that the results simply reflect individual differences across the two studies. Conlon et al. (2001) showed that individuals scoring differently on measures of susceptibility to visual discomfort can show very different spatial frequency tuning for discomfort ratings.

Alternatively, this difference could have been as a result of the slightly eccentric presentation of the stimuli used in the current experiment. The peak of the contrast sensitivity function shifts to lower spatial frequencies as the eccentricity of the stimuli presented is increased (Johnston, 1987). If maximum discomfort occurs for stimuli at the peak of the contrast sensitivity function, a peak at frequencies lower than 3 cycles/degree might be expected, because the centres of both the test and standard stimuli were at an eccentricity of 5.71°. The perceived spatial frequency of gratings is also affected by their eccentricity - a given spatial frequency will be perceived as being a higher frequency when it is presented in the periphery than when it is presented centrally (Davis, Yager, and Jones, 1987). Again, this might be expected to shift the peak of the discomfort function to a lower spatial frequency for peripherally presented stimuli in comparison with centrally viewed stimuli. To address these issues directly, the experiments were repeated with the stimuli presented centrally.

2.5 Experiment 3: Contrast Matching for Centrally Presented Stimuli

The aim of this experiment was to repeat the contrast matching experiment for centrally-presented stimuli. Again, of interest was to establish firstly whether bump stimuli have the same perceived contrast as $1/f$ stimuli that are matched for physical contrast, and secondly whether perceived contrast is affected by the peak spatial frequency of the bump.

Method

Observers: The same two experienced psychophysical observers from Experiment 1 again served as observers.

Stimuli: The stimuli presented were the same as those used in experiment one, except that they were presented centrally.
2.6. EXPERIMENT 4: RELATIVE DISCOMFORT JUDGEMENTS FOR CENTRALLY PRESENTED STIMULI

Procedure: A two-interval forced-choice (2IFC) method was used, in which the test and standard were presented sequentially in the centre of the screen. The first stimulus was presented for 1.8s, with a delay of a single frame before the second stimulus was presented, also for 1.8s. The order of presentation of the stimuli was randomly chosen on each trial. The observers’ task was to decide whether the stimulus in the first or second interval had the higher contrast.

Results
The RMS contrasts required so that each bump stimulus matched the natural slope standard are plotted in Figure 2.8 as a function of spatial frequency. Again, a clear lack of contrast constancy is apparent. Stimuli with bumps centred on 1.5 and 3 cycles/degree again appear significantly higher in contrast than natural slope stimuli matched for RMS contrast, as can be seen from the 95% confidence limits on the estimates of the points of subjective equality. Perceived contrast reduced for higher and lower frequencies. This pattern of results does not differ from that found in experiment one. In particular, there is no shift in the peak of the perceived contrast function to higher spatial frequencies. These results were used to produce stimuli matched for perceived contrast that could be used for a discomfort judgement experiment with centrally-presented stimuli.

2.6 Experiment 4: Relative Discomfort Judgements for Centrally Presented Stimuli
The aim of this experiment was to repeat the second condition of experiment two, using only the 3 cycles/degree bump stimulus standard, for centrally-presented test stimuli. Given the effect of eccentricity on the location of the peak of the contrast sensitivity function (Johnston, 1987), and on perceived spatial frequency (Davis et al., 1987), it might be predicted that the tuning
Relative RMS contrast is plotted (y-axis) against spatial frequency of stimuli (x-axis) for centrally presented stimuli. This again shows a U-shaped tuning function for PSEs for two observers, LOH and PBH, meaning that perceived contrast is not the same across spatial frequencies. Error bars show 95% confidence limits.
of the comfort/discomfort judgements might shift to higher spatial frequencies for centrally-presented stimuli.

Method

Observers: Fifteen naive student observers (eleven female, four male), with normal or corrected-to-normal vision took part in the study. All were within the age range 18 to 30. None of these observers had taken part in the first experiment.

Stimuli: Bump stimuli were the same as in the second condition of experiment two, except that this time they were matched for perceived contrast for centrally-presented stimuli, based on the results of experiment three.

Procedure: A two interval forced-choice (2IFC) method was used, so that the stimuli were presented one after the other centrally. Observers indicated which stimulus they found more comfortable using the arrow keys on the computer keyboard: left for the first interval, right for the second interval.

Results

Results are plotted in Figure 2.9. As in experiment two, all observers showed a similar pattern of results, which were therefore averaged across observers. Spatial frequency tuning remained, and was similar to that observed for the eccentrically presented stimuli in experiment two. Lower spatial frequencies were more often judged more uncomfortable than the 3 cycles/degree bump standard. Again, results of a Friedman’s test showed tuning to be significant: \(\chi^2(6) = 19.97, N = 15, p < 0.05\). Wilcoxon tests again showed that the median responses to spatial frequencies lower than three cycles/degree (0.19 to 1.5 cycles per degree; median = 8.1) differed from those to the three cycles/degree test stimuli (median = -1.5) \((T(14) = 13, p < 0.05)\). The median responses to the two highest frequencies (6 and 12 cycles/degree; median = -3.5) did not differ from those to the 3 cycles/degree test stimuli (median = -1.5); \((T(14) = 39.5, \text{NS})\).
2.7. EXPERIMENT 5: DISCOMFORT RANKINGS FOR PRINTED STIMULI

Figure 2.9: 2IFC discomfort judgements

Percentage of times the test stimulus was chosen as more uncomfortable (y-axis) plotted against spatial frequency of test stimulus (x-axis) for centrally presented stimuli. Error bars show ± 1 standard error.

Discussion

Changing the location of the stimulus to the centre of the visual field did not affect the pattern of results. The spatial frequency tuning was not affected by the central presentation of stimuli, and midrange spatial frequencies were consistently judged more uncomfortable.

The combined results of Experiments 1 to 4 show that there is an effect of spatial frequency on visual discomfort. Even when matched for perceived contrast, stimuli with a concentration of energy around spatial frequencies lower than 3 cycles per degree tended to be judged more uncomfortable than stimuli in which energy is concentrated around 6 and 12 cycles per degree.

2.7 Experiment 5: Discomfort Rankings for Printed Stimuli

A final experiment was performed to address some methodological differences between the experiments, and those performed by Fernandez and Wilkins (2008). In the experiments thus far, stimuli were presented on a computer
2.7. EXPERIMENT 5: DISCOMFORT RANKINGS FOR PRINTED STIMULI

monitor, which was viewed from a fixed distance using a chin-rest. This was to ensure that properties of the images such as their spatial frequency, luminance and contrast were controlled as accurately as possible. A simple 2AFC judgement task was used. Procedures differed from those adopted by Fernandez and Wilkins (2008) in a number of ways. Firstly, they used stimuli that were printed onto paper and held at arms’ length by their participants. Secondly, discomfort was assessed by rating each image, on a 10-point scale, and by sorting them into rank order. A final experiment in the current study, in which observers made rank-order judgements for printed stimuli, was performed to assess the possible effects of these differences on the results.

Method

Stimuli: Stimuli were made as before using the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997) and MATLAB 7.1 (The Mathworks Inc., 2005). They had the same spatial frequencies as those used before (0.188, 0.75, 1.5, 3, 6, 12 cycles/degree), when viewed at a distance of one metre. One metre was marked using a piece of string that was held up to the face whilst viewing the stimuli. Stimuli were matched for perceptual contrast based on the contrast matching data collected in an experiment three. Stimuli were printed so that the visible pattern was 16 x 16 cm on white A4 paper using a HP Laserjet 1200 printer, which was calibrated using the same Minolta LS-110 photometer as previously used. The seven images were presented in a circular formation against a black wall. The centre of each stimulus was at a distance of approximately 10cm from the centre of the circle. The angular extent of the images was 9.1 deg.

Observers: 15 naïve student observers, with normal or corrected-to-normal vision took part in the study. One of these had taken part in Experiments 2 and 3, another four observers had taken part in Experiment 2. All gave informed consent to the study in accordance with the University of St Andrews Teaching and Research Ethics Committee.

Procedure: Observers viewed the set of images, placed in a randomised configuration, and ranked them according to comfort (comfortable to uncomfortable) using the score sheet provided. Viewing time was unlimited.
Results

Average rankings across observers are shown in Figure 2.10. Once again, clear tuning for spatial frequency is evident, with discomfort ratings tending to be lower for the highest spatial frequency present. The shape of the tuning is similar to that obtained in the two previous 2AFC and 2IFC tasks (Experiments 2 and 4). As the data were ordinal (based on rankings), again a non-parametric statistic was used; the results of a Friedman test showed the spatial frequency tuning to be statistically significant ($\chi^2(6) = 12.97$, $N = 15$, $p < 0.05$).
2.8. GENERAL DISCUSSION

Figure 2.10: Printed stimuli discomfort judgements

Average discomfort rankings of printed stimuli (y-axis) plotted against spatial frequency (x-axis). Spatial frequency tuning effects are once again shown. Error bars show standard error.

Discussion

The results of this experiment again show clear effect of spatial frequency on visual discomfort judgements. The results differ somewhat from those found in the earlier experiments, as a marked effect is only evident for the highest spatial frequencies in this case. This difference may reflect the greater variability in the data obtained from rank-order judgements compared to 2AFC tasks. Nevertheless, these results show that overall the effect is robust to differences in terminology used in the instructions (whether participants were asked about comfort or discomfort), to method of presentation of the stimuli (printed vs. presentation on a computer monitor) and method of judgement (2AFC comparisons vs rankings of all stimuli).

2.8 General Discussion

Previous studies have shown clear effects of both the contrast and spatial frequency of stimuli on their judged discomfort (Wilkins et al., 1984; Fernandez and Wilkins, 2008; Juricevic et al., 2010). These effects have been linked both
to deviations of the statistics of stimuli from those that are typical of natural images, and to a role played by hyperexcitation of the visual cortex in discomfort (Juricevic et al., 2010). In the current study, the potential roles of spatial frequency and perceived contrast in these effects were uncoupled. Broadband stimuli of the type used by Fernandez and Wilkins (2008) do not exhibit full contrast constancy. That is, stimuli that are matched for physical contrast do not necessarily have the same perceived contrast. With the potential role of perceived contrast taken into account, it was found that adding energy to a $1/f$ stimulus within a relatively narrow frequency band tended to increase their discomfort. That is, the results of Fernandez and Wilkins (2008) are unlikely to have occurred simply as a result of possible differences in perceived contrast. Moreover, this effect was tuned for spatial frequency, such that the addition of energy around 0.375 to 1.5 cycles/degree of spatial frequency produced the greatest discomfort ratings. These findings were robust to variations in the presentation of the stimuli, or the rating method used. Similar results were obtained regardless of whether stimuli were presented on a computer monitor or printed onto paper, or presented centrally or peripherally. Similar results were also obtained regardless of whether discomfort was judged by rank ordering, or using 2AFC and 2IFC methods.

The tuning of these ratings peaks around frequencies of 0.75 to 1.5 cycles/degree. This is lower than the peak around 3 cycles/degree discussed by Fernandez and Wilkins (2008), although within the range of within two octaves suggested. However, it is important to note this conclusion was based on an analysis of the spectra of comfortable and uncomfortable artworks and photographs, and that the location of the spatial frequency peak was not directly manipulated in their experiments. Earlier work by Wilkins et al. (1984) suggested that the critical midrange frequencies may lie between 2 and 8 cycles/degree.

Importantly, Conlon et al. (2001) showed clear individual differences in these results, dependent on participants’ susceptibility to visual discomfort. Since this was not reported by Fernandez and Wilkins (2008), and was not measured in the current study, this might be one factor that could contribute to differences in the exact range of frequencies responsible for discomfort in different studies.

As discussed in the introduction, a number of explanations of the discomfort evoked by certain stimuli have been proposed. Firstly, discomfort may reflect hyperexcitation of the visual cortex. This would then explain why it is those frequencies to which the visual system is most responsive that are typically as-
associated with discomfort. The results of the current study do not conflict with this idea. However, any such effects might be expected to vary with perceived contrast, and therefore should be minimised in the experiments since in all cases stimuli were approximately matched for perceived contrast. A second, related explanation is that discomfort arises when the statistical properties of images deviate significantly from those of natural images (Juricevic et al., 2010). Since it is reasonable to assume that the visual system is optimised to encode natural images, such deviations will lead to suboptimal encoding. Specifically, Field (1994) noted that the spatial frequency bandwidths of visual channels mean that they will carry equal amounts of information in natural images. When energy is concentrated around a narrow range of spatial frequencies, the distribution of responses across channels will therefore be uneven. This departure from the expected distribution of activity might be responsible for the discomfort. This might occur through a change in the sparseness of the response (Redies et al., 2007), or to the metabolic load placed on the system (Barlow, 1961). In attempting to quantify such possible differences, it will also be necessary to take account of the time-course of the visual system’s responses, and how this would be affected by temporal properties of the stimuli. Detailed consideration of these issues is beyond the scope of the current study and will be discussed in Chapter 3.

A final possible factor is differences in the perceived blur, and accommodative load, across stimuli. Accommodative insufficiency, and accommodative lag (in which the image is focussed in front of the retina), are associated with visual discomfort (Chase et al., 2009; Allen, Hussein, Usherwood, and Wilkins, 2010). Accommodative lag can occur after reading for prolonged periods of time, possibly through fatigue of the system due to high accommodative demand. One cue that drives accommodation is the perceived blur of the stimulus (Horwood and Riddell, 2008), which in turn is affected by its amplitude spectrum. A concentration of energy at relatively low spatial frequencies occurs when a stimulus is not well-focussed, and will cause it to appear blurred (Murray and Bex, 2010; Webster, Georgeson, and Webster, 2002). It is possible therefore that some of the changes in the amplitude spectrum in the current study, and that of Juricevic et al. (2010) are interpreted as blur caused by poor accommodation. This blur might then act as a stimulus for accommodation (Day et al., 2009a), and the fact that this blur cannot be removed by changes in accommodation might lead to discomfort, possibly through the fatigue of the accommodation system.

Given the importance of accommodative insufficiency as a determinant of
visual discomfort (Chase et al., 2009; Allen et al., 2010), measures of the refractive error and accommodative responses of individuals might be useful when accounting for how it is affected by the spatial properties of images. However, such an analysis of individual differences in responses to stimuli goes beyond the scope of the current study, which aimed to assess some of the spatial factors affecting discomfort in the general public. The appropriateness of stimuli for driving the accommodative response will be investigated in more detail in Chapter 4.

Overall, a detailed, quantitative understanding of the role of contrast and spatial frequency on visual discomfort is important to help to consolidate what is already known from clinical literature and the characteristics of uncomfortable images. This will be a vital contribution to a better understanding of the functioning of the visual system, and how to avoid visual discomfort in normal and clinical populations. For example, almost all work concerned with the optimal encoding of visual information has addressed how this might be achieved, rather than the consequences of less than optimal responses. Studying the causes of visual discomfort is also important in order to understand and prevent adverse effects for clinical populations such as migraineurs and epilepsy sufferers, as well as to provide benefits for non-clinical populations such as increased productivity (e.g. Daum, Clore, Simms, Vesely, Wilczek, Spittle, and Good, 2004) and better reading ability, as well as guidelines of how uncomfortable stimuli can be avoided in the media (Wilkins, Emmett, and Harding, 2005).

As previously discussed, one possible reason why these spatial frequency manipulations might cause discomfort is through inefficient coding. Sparse coding models have been used to show how natural images are processed efficiently. Sparse coding depends on the spatial frequency content of the image: in order for images to be optimally processed by the visual system, they must have the spatial frequency characteristics that allow this. It might be the case than spatial frequency manipulated stimuli are uncomfortable as they do not allow for efficient coding. How efficiently these stimuli are encoded will be investigated in Chapter 3, using a sparse coding model based on that of Field (1994).
CHAPTER 3
MODELLING SPARSE CODING

3.1 Introduction

From the results of the experiments in Chapter 2, it is clear that the spatial frequency content of images affects discomfort judgements. These effects could be due to the nature of the neural responses to those images. It has been suggested by Juricevic et al. (2010) that one possible reason for discomfort from visual stimuli is overly large responses in the brain. The aim of this chapter is to investigate this possibility.

It is possible that uncomfortable stimuli do not facilitate efficient coding in the brain, which might result in such a large cortical response. It requires significant metabolic energy to process visual information in the brain (Lennie, 2003). Efficient coding of visual stimuli would rely on maximising information transfer whilst minimising metabolic cost (Levy and Baxter, 1999). This could potentially be achieved by exploiting the statistics of the image to reduce redundancy through the encoding process (Barlow, 1961). Natural images are typically redundant - there is information contained in them that could be extrapolated by knowing other bits of information about the image. By contrast, random white noise images are not redundant - there is no way of knowing what value a particular pixel will have from knowing adjacent pixel values. Under normal circumstances, the visual system would tend to encounter mostly natural images, which have particular statistics, such as the $1/f^\beta$ amplitude spectrum (Tolhurst et al., 1992). It has been suggested that these predictable statistics allow natural images to be processed efficiently.

Sparse coding models present a theoretical account of how the visual system
3.1. INTRODUCTION

might efficiently process natural images (Field, 1987, 1994, 1999; Simoncelli and Olshausen, 2001). These models are based on the principle of having only a very few neurons respond strongly to a stimulus, and the majority of neurons responding only weakly. Using a simplified model of the visual system, it is possible to predict how efficiently a stimulus could be processed by the visual system (e.g. Field, 1994). Images that are processed efficiently result in highly kurtotic responses: most units do not respond above baseline, whilst a few units respond strongly. In this manner it is thought that maximum information can be transmitted with minimal metabolic cost.

The Visual System

After the retina, the main pathway for visual information is through the geniculo-cortical pathway, which is composed of the lateral geniculate nucleus (LGN) in the thalamus, and the primary visual cortex (V1) in the occipital lobe (see DeValois and DeValois, 1990, for a detailed discussion). It has been suggested that the LGN is responsible for the initial band-pass filtering of stimuli in the cat LGN (Kaplan, Purpura, and Shapley, 1987; Tan and Yao, 2009). This helps maximise information transfer by reducing redundancy in low signal to noise ratios by filtering out the high spatial frequency content likely to be noise, and reducing the input of the low spatial frequencies, which are typically high amplitude (Kaplan et al., 1987). Something analogous in humans could lead to the different contrast sensitivities for different spatial frequencies - people are more sensitive to midrange spatial frequencies that either high or low spatial frequencies (Campbell and Robson, 1968). The tuning of the LGN might contribute to this by reducing the signal of the lower spatial frequencies, and filtering to remove the high spatial frequencies.

Hubel and Wiesel (1962) showed that cells in the primary visual cortex are both spatial frequency and orientation tuned. DeValois et al. (1982) suggested that cortical cells are more narrowly tuned for spatial frequency and orientation than LGN cells. The authors argued that while LGN simple cells respond to a range of spatial frequencies and orientations, cortical cells respond very specifically to stimuli of their own preferred phase, orientation and spatial frequency. It has been suggested that the LGN is mainly involved in preprocessing the image to remove noise, hence the tuning to remove high spatial frequency content, which is likely to be noise, as anything over around 50 cycles/degree will be filtered out by optics (Campbell and Green, 1965). By contrast, the role of the cortex might be to extract the relevant information from the image, and thus various filters specific for different spatial frequencies, orientations
3.1. INTRODUCTION

and phases extract the information from the preprocessed signal. Evidence from DeValois et al. (1982) shows that there is a relatively continuous range of spatial frequencies that cortical cells respond to.

The responses of these two main stages of the visual system to images have been modelled in the past using wavelet filters (e.g Field, 1994). The Field (1994) model used convolutions of an input image by arrays of mathematically defined model ‘cells’ to determine what the response of two different levels of visual processing should be. Field’s model cells were based on some physiological characteristics of the cells found in these two areas. The LGN was modelled with a difference-of-Gaussian (DOG) filter, to recreate the centre-surround properties of these cells (see Figure 3.1 for an example of the LGN Difference-of Gaussian filter). DOG filters have the same properties as the LGN cells, based on biological evidence (Kaplan, Marcus, and So, 1978). V1 cells were modelled with more elongated wavelet filters, known as ‘Gabor’ filters. Gabor filters are chosen for the cortex as their properties show a good match to the tuning functions of simple cells (see Daugman, 1988). These also had ‘ON’ and ’OFF’ areas, but were selective for orientation and phase also. See Figures 3.2 and 3.3 for examples of the filters.

Measuring Sparseness

In his model, Field (1994) showed there to be sparse coding resulting from the responses of the cells. He measured ‘sparseness’ in terms of response kurtosis. Kurtosis is the fourth moment of a distribution. It is defined by Equation 3.1:

\[
Kurtosis = \frac{\mu^4}{\sigma^4}
\]  

(3.1)

Kurtosis: this is the measure of the fourth moment about the mean. \( \mu \) is sum of the the difference between the score and the mean, \( \sigma \) is the standard deviation of the distribution. Kurtosis of a Gaussian (normal) distribution is three, therefore ‘excess kurtosis’ is also sometimes calculated, by subtracting
3.1. INTRODUCTION

Figure 3.1: Difference of Gaussian filter

A cross-section of the DOG filter representing the LGN.

Figure 3.2: Gabor filter

A cross-section of the Gabor filter representing V1.
3.2. THE MODEL

Examples of Gabor filters at various spatial frequencies, orientations and phases. This kind of filter makes up the model cortex.

Field (1994) found there to be higher kurtosis with later stages of processing of natural images: the LGN cells increased kurtosis from the original image, and V1 increased kurtosis even further. Thus Field concluded that the goal of sensory processing is to increase sparseness in the response. This is theoretically important; increasing sparseness would allow maximum information to be transmitted with minimal metabolic expenditure in terms of populations of neurons firing. This is of importance to visual discomfort as it is thought that visual discomfort might arise from excessive cortical activity (Juricevic et al., 2010).

The goal of this current model is to represent the response of a population of different orientation and spatial frequency tuned cells, each cell type represented by a different filter. The population response is of interest here, given the nature of the sparse coding hypotheses. In this current model, the ‘cells’ of the model will be based on physiology taken from experimental data such as single cell recordings (e.g. DeValois et al., 1982).

3.2 The Model

Field (1994) modelled the LGN as two difference-of-Gaussian filters, one centred on 20 cycles per image, the other at 40 cycles per image. See Equation 3.2.
3.2. THE MODEL

\[ g(x, y) = 9 \exp\left(-\frac{r^2}{2(3\sigma_1^2)}\right) - \exp\left(-\frac{r^2}{2\sigma_2^2}\right) \]  

(3.2)

Difference of Gaussian (DOG) filter. There were two filters: \( \sigma \) was defined to result in either 0.2 and 0.4 degrees width at half height, \( r \) was the radius.

The peak responses of the LGN filter model were adjusted, based on physiological data from single cell recordings in the cat (Tan and Yao, 2009), and in the cat and macaque (Kaplan et al., 1987) to have a peak response at around 0.2 cycles per degree and 0.4 cycles per degree. It is thought that the LGN acts as a band pass filter (Tan and Yao, 2009), and therefore this might have effects such as increasing kurtosis from that of the original image. In order to see the effects of the LGN component on the response, this will be examined separately, as well as examining the combined response from the LGN and the cortex.

Field (1994) modelled V1 simple cells in the cortex as a wavelet filter, a log-Gabor, based on the work of Field (1987). In the papers by Field (1987) and Field (1994), a log-Gabor was used, to avoid overrepresentation of higher spatial frequencies. The current model uses a Gabor model of cortical V1 cells, see Equation 3.3. A Gabor has been shown to be a good approximation of the receptive fields of V1 cells (Daugman, 1989).

\[ G(k) = s_1 \exp\left(-\frac{x_p^2}{2\sigma_1^2} - \frac{y_p^2}{2\sigma_2^2}\right) \]  

(3.3)

The Gabor function used in the current model. This consists of a Gaussian modulated by a cosine. Where \( f \) is the spatial frequency, and \( \phi \) is the phase of the Gabor, \( \sigma_1 = 0.39/f \), and \( \sigma_2 = 0.78/f \), after the work by (Hibbard, 2008). \( s_1 \) is a sine wave defined below. \( y_p \) and \( x_p \) are co-ordinates and are defined below. \( A \) is a constant.

\[ s_1 = A \cos(2\pi fx_p + \phi) \]  

(3.4)
3.2. **THE MODEL**

\[ x_p = x \cos \theta - y \sin \theta \]  
\[ (3.5) \]

\[ y_p = y \cos \theta + x \sin \theta \]  
\[ (3.6) \]

The model by Field (1994) model included two spatial frequencies; 20 and 40 cycles per image, the present model expanded on this by including six spatial frequencies. As behavioural evidence indicates that there is an effect of spatial frequency on visual discomfort, it was felt it might be useful to place more emphasis on this in the model. The following spatial frequencies were chosen; 0.5, 1, 1.5, 2, 4, and 12 cycles/degree. The lowest spatial frequency was 0.5 cycles per image due to computational limitations, and the others were chosen based on the physiological data of DeValois et al. (1982). Orientations chosen were 0, 45, 90, 135 degrees, and phase was either odd or even. This resulted in a range of 48 filters to represent the cortex.

**Stimuli:** There were six images of each class. The size of all images was 1201 x 1201 pixels. All images were normalised so that contrast varied between 0 and 1 before analysis. Natural stimuli consisted of six calibrated images of woodland and beach scenes taken from Hibbard (2008). 1 pixel was defined as 1 arcmin in the model.

Two types of filtered noise stimuli were created. The ‘bump’ stimuli as in Chapter 2 were made. This involved filtering random noise in frequency space using a raised radial cosine, see Equation 3.7. Where \( T \) is 0.9, \( \beta \) is the roll-off factor of 0.5, \( f \) is the spatial frequency, and \( f_0 \) is the centre frequency of the peak. The peak frequency of the raised radial cosine was at 0.188, 0.375, 0.75, 1.5, 3, 6 cycles/degree. See Figure 3.4 for examples.
3.2. THE MODEL

Figure 3.4: ‘Bump’ stimuli.

Figure 3.5: Filtered noise stimuli of varying slope exponent, these are 1.5, 2, 2.5 exponent values.

\[
H(f) = \begin{cases} 
    T & \text{for } (0 \leq |\log(f) - \log(f_0)| \leq \frac{1-\beta}{2T}) \\
    \frac{T}{2} \left[ 1 + \cos\left( \frac{\pi T}{2} \right) \right] \left( |\log(f) - \log(f_0)| - \frac{1-\beta}{2T} \right) & \text{for } \left( \frac{1-\beta}{2T} \leq |\log(f) - \log(f_0)| \leq \frac{1-\beta}{2T} \right) \\
    O & \text{for } (|\log(f) - \log(f_0)| > \frac{1-\beta}{2T}) 
\end{cases}
\]  

(3.7)

A second type of filtered noise pattern used was $1/f^\beta$ filtered noise with the exponent of the slope ($\beta$) manipulated. $\beta$ was either 0.5, 0.75, 1, 1.5, 2, 2.5. Again these stimuli were made by filtering random noise in frequency space. See Figure 3.5 for examples.
3.2. **THE MODEL**

Figure 3.6: Sine-wave striped stimuli.

Striped patterns were created using MATLAB (The Mathworks, Inc, 2005) and the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997; Kleiner, Brainard, and Pelli, 2007). Striped patterns included sine waveforms of differing spatial frequency were created using Equation 3.8. Spatial frequency of these waveforms was varied: 0.375, 0.75, 1.5, 3, 6, 12 cycles/degree, corresponding to 7.5, 15, 30, 60, 120, and 240 cycles/image. See Figure 8.4 for examples. Square waveforms of varying spatial frequency were also included: again 0.375, 0.75, 1.5, 3, 6, 12 cycles/degree, corresponding to 7.5, 15, 30, 60, 120, and 240 cycles/image. See Figure 3.7.

\[ y = A \sin (2\pi f \theta) + \phi \]  

(3.8)

Where \( A \) is a constant, \( x \) and \( \theta \) create the horizontal displacement, \( f \) is the frequency of the stripes, and \( \phi \) is the orientation of the stripes.

The work of Bridget Riley has been investigated by Zanker (2004); Zanker, Herman, and Walker (2010). Although they were not interested in discomfort *per se*, they were investigating the possibility that illusions reported on viewing these stimuli are associated with very small, fast eye movements called microsaccades. These shimmering illusions are of interest as they come under the remit of ‘discomfort’ according to other authors (e.g. Wilkins and Evans, 2001), and the work of Bridget Riley has been studied in previous work on
discomfort (e.g. Fernandez and Wilkins, 2008). In these papers, Zanker (2004) and Zanker et al. (2010) used one of Riley’s most famous artworks ‘Fall’ to create stimuli (‘riloids’) that were thought to induce inconsistent local motion signals. The stimuli for the next part of the analysis were created by using their ‘riloid’ equation in order to systematically vary the spatial frequency of the basic waveform and investigate these effects on the model. See Equation 3.8.

\[
I(x, y) = 0.5 \left(1 + c \sin \frac{2\pi(x - \phi y)}{\lambda}\right)
\]

(3.9)

Where \( c \) is the contrast, \( \lambda \) is the wavelength and \( \phi \) is the spatial phase of the grating and is given by

\[
\phi(y) = A \sin 2\pi \frac{y}{\mu y}
\]

(3.10)

The phase modulation \( \mu \) was fixed and therefore constant all the way down the riloid pattern at 180 pixels. The wavelength is defined as \( 1/f \), and the same set of spatial frequencies were used as in the stripe stimuli; 0.375, 0.75, 1.5, 3, 6, 12 cycles/degree, translating to 7.5, 15, 30, 60, 120, 240 cycles/image. See Figure 3.8 for examples.
Op-art images were taken as GIF files from www.viperlib.york.ac.uk. These were converted to greyscale JPEG format and resized to 1200 x 1200 pixels using the Matlab (The Mathworks, 2005) functions ‘im2double’ and ‘imresize’.
3.2. THE MODEL

Figure 3.9: ‘Blue Sun’ Image supplied by Akiyoshi Kitaoka at the College of Letters, Ritsumeikan University, Kyoto (Japan).
Figure 3.10: ‘Clarch01’ Image supplied by Akiyoshi Kitaoka at the College of Letters, Ritsumeikan University, Kyoto (Japan).
3.2. THE MODEL

Figure 3.11: ‘Enigma 780’ Image supplied by Michael Bach at the Augenklinik, University of Freiburg, Germany. Image originally created by Isia Leviant (1984).
Figure 3.12: ‘Gira 3’ Image supplied by Akiyoshi Kitaoka at the College of Letters, Ritsumeikan University, Kyoto (Japan).
Figure 3.13: ‘Warp’ Image supplied by Akiyoshi Kitaoka at the College of Letters, Ritsumeikan University, Kyoto (Japan).
3.2. THE MODEL

Figure 3.14: ‘Warp 7’ Image supplied by Akiyoshi Kitaoka at the College of Letters, Ritsumeikan University, Kyoto (Japan).
3.3. RESULTS

**Analysis:** Sparseness of response was the measure of interest. Following Field (1994), kurtosis was taken as the measure of the sparseness of the response. Kurtosis is defined as Equation 3.1. The kurtosis of the image was calculated by using the inbuilt function ‘kurtosis’ in MATLAB. The DOG filter responses were turned into a single vector, and the ‘kurtosis’ function was used to calculate the kurtosis of these. All 48 cortical filter outputs were vectorised and the kurtosis of these was also taken using this function.

3.3 Results

**Kurtosis Increases With Processing**

Field (1994) demonstrated that response kurtosis increases with subsequent processing by the model visual system. The model output confirms this; response kurtosis was found to increase with more levels of processing in the model visual system. This can be seen in Figure 3.15. Kurtosis of the original images is relatively low compared with the model output at later stages of processing.

![Model Responses](image)

Figure 3.15: Response kurtosis with increasing level of processing: original image kurtosis, DOG (LGN) and Gabor (cortex) filters.
3.3. RESULTS

This is not unexpected: filtering images will in general increase kurtosis, as some responses are reduced due to the filtering process. What is of more interest to the current investigation is whether there is a difference in sparseness between the responses from comfortable and uncomfortable images. Field (1994) places much emphasis on kurtosis as a measure of sparseness of responses to natural images. Therefore this metric was used to investigate whether differences in response kurtosis were evident that could potentially account for reported discomfort.

Comparing Response Kurtosis

Mean response kurtosis from different image series are plotted separately for the different levels of processing in Figures 3.16 (LGN) and 3.17 (cortex). Note the difference in axes - the kurtosis is much higher at the model cortex level compared to the model LGN. From this plot, it can be seen that the average response kurtosis for natural images is higher than for any of the uncomfortable images. This is both at the level of the LGN (DOG filter) and also at the level of the cortex (Gabor filter). This is important as it means the response distribution is more peaked. This could indicate that many units are responding strongly, but most are not.

Response kurtosis was an emphasised component of sparse coding models (e.g. Field, 1994). A high response kurtosis could mean that few neurons respond strongly, but others do not respond much. However, it could also be a highly kurtotic response if many neurons respond strongly and a few are inactive. This would appear as a U-shaped distribution. Therefore it is also critical that the mode response is zero. This means that most of the units are inactive, with a strong response from only a few. It is possible to have high kurtosis with most units responding strongly, which would not be a sparse encoding. Thus mode responses are now considered.

Mode Responses

Mode response from each image was calculated for both output levels of the model. The mean of the mode responses for all six images within a class
3.3. RESULTS

Figure 3.16: Response kurtosis

Response kurtosis for the model DOG (LGN) filters. Nat=natural, str=stripes, sst=square-wave stripes, ril=riloids, op=op-art, fno=filtered noise, bum=bumps.
3.3. RESULTS

Figure 3.17: Response kurtosis

Response kurtosis for the model Gabor (cortex) filters. Nat=natural, str=stripes, sst=square-wave stripes, ril=riloids, op=op-art, fno=filtered noise, bum=bumps.
was then calculated. This average mode response for all six images for the model LGN is plotted in Figure 3.18. The average mode response (mean of the mode response for all six images) for the model cortex is plotted in Figure 3.19. These figures were plotted with the same y-axis to show the mode model LGN responses are larger than the mode model cortex responses. The mode response of these images seems to be high at the level of the LGN, but reduces dramatically at the level of the cortex. The response output of the response itself is to some extent arbitrary, it is the distribution of the response that is of interest here. As there is wide variation in the response output of the model to different images, presenting the mode itself might not be as informative as it could be. For example, two identical mode values might be interpreted differently if the maximum response is low, compared to if the maximum response is high. If the maximum response is low, then the mode response would be a much larger percentage of the overall value compared to the high overall response, therefore would potentially indicate more sparse coding in the latter case, despite identical mode values. Therefore the mode responses were also normalised to present the mode response of the filters, relative to the maximum response for that image. This was achieved by dividing the mode response by the maximum response for that image. This normalised mode response can be seen in Figures 3.20 and 3.21, for the LGN and cortical levels respectively.

The normalised mode responses of the LGN show a large difference between some of the uncomfortable stimuli and natural images: natural images, filtered noise and bump stimuli have a low mode response, and the striped patterns seem to have a higher mode response, compared to the other images. However, it is possible that this is due to the sampling. As the stripes are a regular pattern, it might be the case that the model LGN cells might match the spatial frequency of some of the stripes, thus creating an artificially high mode response. Therefore caution is necessary for interpreting this output. Both the mode response itself, and the normalised mode responses, show that by the level of the cortex output, all mode responses are very much reduced. As the normalised mode cortical response of many of the images was low, this indicates that most units are not responding strongly in comparison to the maximum output of the model to that image. Kurtosis and mode are methods for summarising the responses. Thus a more detailed assessment of each distribution individually would be of benefit here.
Figure 3.18: Mode response

Mode response for the various stimulus classes for the DOG (LGN) responses. Nat=natural, str=stripes, sst=square-wave stripes, ril=riloids, op=op-art, fno=filtered noise, bum=bumps.
3.3. RESULTS

Figure 3.19: Mode response

Mode response for the various stimulus classes for the Gabor (cortex) filter responses. Nat=natural, str=stripes, sst=square-wave stripes, ril=riloids, op=op-art, fno=filtered noise, bum=bumps.
3.3. RESULTS

Figure 3.20: Mode response

Normalised average mode response for the various stimulus classes for the DOG (LGN) responses. Nat=natural, str=stripes, sst=square-wave stripes, ril=riloids, op=op-art, fno=filtered noise, bum=bumps.

Response Distribution

Figures 3.22 and 3.23 shows the response distributions from the Gabor (cortical) output for the natural, bump and filtered noise stimuli. All six of the natural images have been averaged and plotted as the solid black line. From this plot it can be seen that the average of the six natural images shows a similar type of characteristic sparse response: a peaky distribution centred around zero. The bump and filtered noise stimuli both show similar distributions to the natural images, despite systematic variation of the relative spatial frequency content in these stimuli.
3.3. RESULTS

Figure 3.21: Mode response

Normalised average mode response for the various stimulus classes for the Gabor (cortex) filter responses. Nat=natural, str=stripes, sst=square-wave stripes, ril=riloids, op=op-art, fno=filtered noise, bum=bumps.
Figures 3.24 and 3.25 show again the average response distribution from the six natural images as the solid black line for comparison. The response distributions for each of the six spatial frequencies of the sine and square waves are shown in this figure. The responses of the model to low spatial frequency (0.375 and 0.75 cycles/degree) sine and square waves shows there to be a much broader distribution compared to the higher spatial frequencies. The model response to the higher spatial frequencies, 3, 6, and 12 cycles/degree, are considerably narrower distributions than those of the natural images. It would appear that the striped patterns are more sparsely distributed than the natural images, as there is a more peaked response distribution. This will be discussed later.

Figures 3.26 and 3.27 show again the average response distribution from the six natural images as the solid black line for comparison. Again, note that the scale is not the same for all plots. Figures 3.26 and 3.27 shows the distributions of responses for the riloid patterns of increasing spatial frequency, and the six op-art images. The op-art images typically have a wider distribution than the natural images. The riloid patterns seem to have distributions similar to those of natural images. The op-art results appear to be less sparse than the natural images, but again the riloids do not support the idea that discomfort might arise from a lack of sparse coding.

3.4 Discussion

The results of the model have replicated the findings of Field (1994) in that response kurtosis increased with each subsequent level of processing. This was the case for all image categories. This can be accounted for in terms of the
model properties alone: filtering could have the effect of reducing the output, as some content is filtered out. Therefore filtering could potentially change the shape of the distribution by reducing the numbers of extreme values - a larger proportion of the distribution of responses will have the value zero. If this is the case, it is perhaps unsurprising that the more levels of processing, the more filtering occurs, and the more kurtotic the response. This has been identified as a potential limitation for such filter-based models (Baddeley, 1996). However, this is not necessarily the case, as the filtering may also affect the mean and other moments of the distribution. For example, in the case of notch filtering, kurtosis may remain unaffected whilst the shape of the distribution is dramatically changed. This emphasises the importance of assessing the entire distribution, rather than simply reducing it to a summary value.

Response kurtosis is emphasised as an important characteristic of sparse codes (Field, 1994). If uncomfortable images are inefficiently processed, it is expected that response kurtosis might show variation between the uncomfortable compared to natural images. Investigating the kurtosis of the different image categories showed that there was no qualitative difference between the natural and uncomfortable images: all image categories increased in kurtosis to a similar degree from the model LGN (DOG) to the model cortex (Gabor) outputs. However, there was a difference in terms of degree: response kurtosis was highest overall for natural images compared to any of the other images.

Kurtosis is not the only necessary aspect of sparse coding: it is also critical that the mode response is zero, indicating most units do not respond. Mode responses of the LGN level showed a difference between the size of the mode response for the natural, filtered noise and bump stimuli compared to the other uncomfortable images. However, mode responses of the cortex were very much reduced compared to the LGN, indicating that the kurtosis can be used as a measure of sparseness of response.

By investigating the overall distribution of the model cortex output it was clear that natural images and both types of filtered noise pattern showed similar types of distributions: kurtotic, with a peak at zero. Effects of spatial frequency manipulations were not obvious for the filtered noise patterns, which generally tended to have similar response distributions as natural images.

Op-art stimuli showed broader distributions compared to natural images. These were much less peaked than some of the other image classes. This is not a sparse code, compared to the other image classes - many units are responding
strongly, and relatively fewer are inactive. This might indicate more metabolic activity, as a consequence of the more distributed response. This is in agreement with the notion that some types of discomfort might be due to inefficient neural processing. If these stimuli do not facilitate a sparse response, this might cause the discomfort. This was not apparent from the kurtosis and the mode, suggesting that these measures alone might not capture enough of a distribution to really quantify sparseness, and additional measures are needed (Baddeley, 1996).

The striped patterns typically showed different distribution shapes compared to natural images. In particular, striped patterns were particularly influenced by spatial frequency. Critically, the striped patterns showed a more sparse distribution than natural images, as the response was more peaked. In addition to this, Riloid and square-wave stimuli showed response distributions quite similar to those of natural images. This is unexpected for this kind of image, as these images could be considered in a similar category as the striped patterns, but their spatial properties do not reflect this. None of these response distributions are in accordance with the idea that discomfort might arise from a lack of sparse coding, as they are either equally, or more sparsely distributed than the natural images.

It is also possible that the discomfort might still arise from an excessive response, but possibly one that is restricted to a single channel. There are inhibitory processes between the spatial frequency channels (e.g. Klein, 1980). When an input stimulates more than one spatial frequency channel, by consisting of a range of spatial frequency information, there is inhibition between them. If the stimulus activates one channel exclusively, it might be the case that this inter-channel inhibition is reduced, which might increase the overall output. The current model does not contain inhibitory processes, therefore this possibility is left for future work.

Alternatively, there might be a completely different source of discomfort in the case of these images, which is equally unable to be accounted for in the current model. For example eye movements have been proposed as a source of discomfort from such periodic stimuli (Troncoso, Macknick, Otero-Millian, and Martinez-Conde, 2008; Zanker, 2004; Zanker et al., 2010). It is also possible that periodic stimuli cause difficulty for binocular co-ordination. The effects of text regularity will be investigated in Chapter 7, using reading as a task.

Additionally, the filtered noise patterns also showed similar shaped distribu-
tions to those of natural images. This suggests that discomfort from filtered noise images might not be accounted for by non-sparse neural responses. This could not be accounted for easily in terms of eye movements in the same way as the striped patterns (e.g. Zanker, 2004; Zanker et al., 2010), as these stimuli are not periodic. It has been shown that the lower spatial frequency manifestations of these stimuli are more likely to be judged as uncomfortable (see Chapter 2 and Juricevic et al. (2010)). Images consisting of high levels of low spatial frequency information are perceived as blurred (Murray and Bex, 2010). Blur is also associated with accommodative response control (e.g. Charman and Tucker, 1977; Charman, 1979). Therefore another possibility is that discomfort from noise patterns might result from a different source to that of striped images, and this could be associated with image blur and accommodative responses. This possibility will be investigated in the next chapter.
3.4. DISCUSSION

Figure 3.22: Cortical response distributions

Model cortical response distributions for each of the six natural images and filtered noise images. Average natural image distribution is plotted on each graph in black solid lines. Bump stimuli are from 0.188, 0.375 and 0.75 cycles/degree. Slope manipulations (fnoise) are lowest to highest slope exponent, 0.5, 0.75 and 1.
3.4. DISCUSSION

Figure 3.23: Cortical response distributions

Model cortical response distributions for each of the six natural images and filtered noise images. Average natural image distribution is plotted on each graph in black solid lines. Bump stimuli are from 1.5, 3 and 6 cycles/degree. Slope manipulations (fnoise) are lowest to highest slope exponent, 1.5, 2, and 2.5.
Figure 3.24: Cortical response distributions

Model cortical response distributions for each of the six sine and squarewave spatial frequency manipulations. Sine and square waves increase in spatial frequency, 0.375, 0.75 and 1.5 cycles/degree. Average natural image distribution is plotted on each graph in black solid lines for comparison.
3.4. DISCUSSION

Figure 3.25: Cortical response distributions

Model cortical response distributions for each of the six sine and squarewave spatial frequency manipulations. Sine and square waves increase in spatial frequency, 3, 6 and 12 cycles/degree. Average natural image distribution is plotted on each graph in black solid lines for comparison.
3.4. DISCUSSION

Figure 3.26: Cortical response distributions

Model cortical response distributions for each of the riloids, and op-art images. Spatial frequency of riloids increases: (0.375, 0.75 and 1.5 cycles/degree). Average natural image distribution is plotted on each graph in black solid lines for comparison.
3.4. DISCUSSION

Figure 3.27: Cortical response distributions

Model cortical response distributions for each of the riloids, and op-art images. Spatial frequency of riloids increases: (3, 6 and 12 cycles/degree). Average natural image distribution is plotted on each graph in black solid lines for comparison.
CHAPTER 4

BLUR AND VISUAL DISCOMFORT

4.1 Introduction

From the results of Chapter 2, it has been shown that manipulating the spatial frequency content of filtered noise affects discomfort judgements. In particular, images with a relatively higher contrast amplitude at low spatial frequencies compared with the statistics typical of natural images were judged more uncomfortable than stimuli with relatively increased contrast amplitude at high spatial frequencies. This is in agreement with other studies showing that increasing the relative amplitude at low spatial frequencies increased discomfort judgements (Juricevic et al., 2010).

Relatively less high, and more low spatial frequency information, is a characteristic often associated with perceived blur (e.g. Murray and Bex, 2010). The spatial frequency content of the image provides an important cue for the accommodation (focussing) response (e.g. Charman and Tucker, 1977; Charman, 1979; MacKenzie et al., 2010). The aim of this study was to investigate the relationship between excess low spatial frequency information, discomfort judgements and perceived blur. This was investigated using relative discomfort judgements in both artificial and natural images. A possible explanation for this relationship relating to accommodation responses is proposed. This potential explanation was investigated by modelling the retinal image resulting from comfortable and uncomfortable stimuli. The model suggests that the stimuli identified as more uncomfortable from the psychophysical tasks might present a less informative stimulus to drive the accommodative response than others.
4.1. INTRODUCTION

Blur

An increase in low spatial frequency information could potentially be interpreted as blur, as it would be the result of low-pass filtering the image. However there are many different ways of quantifying blur. Global luminance statistics are one measurement of blur, others include the steepness of the local contrast gradients in the image.

This can be illustrated by considering a square-wave (step edge) compared to a sine wave luminance gradient. One way to characterise the image is the amplitude spectrum, which relates to the relative amount of spatial frequency information available from the whole image at different spatial scales. In a step edge, the amplitude of the harmonics falls with a $1/f$ pattern: higher harmonics have smaller amplitudes. Removing harmonics from this waveform has the effect of decreasing the high spatial frequency content, making the image more like a sine wave.

The amplitude of the information at the different spatial frequencies, the $1/f$ relationship, is not the only contributing factor to blur perception. Phase manipulations affecting the luminance gradients of waveforms also affect the perceived blur of an image (e.g. Badcock, 1984). A sharp edge has a step luminance profile; a steep transition from black to white and vice versa. The blurred sine wave edge has a shallower luminance profile, or contrast gradient. The more blurred, the shallower the luminance gradient contrast gradients. The characteristics of the image in the spatial domain are also important to whether the image is seen as blurred. The relative location of the spatial frequency information depends on the phase of the information contained at each spatial frequency, known as the image phase spectrum. This relates to the phase component of the Fourier transform of the image, which contains information about the relative phases of the components of the image. As an example, this can be thought of as the difference between a sharp, step edge (all scales of information in phase) compared to a smoother decline of contrast (out of phase). Thus images can have the same amplitude spectrum, but a shallower luminance gradient, if the component spatial frequencies are out of phase compared to in-phase.

The steepness of the luminance gradients can be manipulated by reducing the amplitude at high spatial frequencies, thus increasing the slope of the amplitude spectrum. Alternatively luminance gradients can be manipulated by changing the phase relationships of the information in the image (e.g. Bad-
4.1. INTRODUCTION

cock, 1984). Therefore an image could be perceived as blurred because of loss of high spatial frequency information overall, or alternatively might be due to local contrast gradients, determined partly by phase relationships. Both manipulations would affect contrast gradients and possibly be perceived as blur. If blur more generally is the cause of discomfort, it might be expected that blur induced by phase as well as amplitude manipulations also affects discomfort judgements. Whilst the distinction is straightforward in artificial stimuli, this becomes more complex when considering blur in natural images.

![Figure 4.1: Amplitude spectra of natural images.](image)

Top row shows different image manipulations, the bottom row shows the corresponding amplitude spectra. Luminance amplitude is plotted against log spatial frequency (cycles/image). Points on the graph are individual pixels from the Fourier amplitude spectrum. These indicate contrast amplitude at each spatial frequency, sampled at all orientations. Figure A shows the original image and its amplitude spectrum. Figure B shows an image in which the slope exponent ($\beta$) of the original image (1.45) has now been increased to 2. This steeper slope results in increased amplitude at the lower spatial frequencies. Figure C shows a Gaussian blurred version of the original image. Again, the amplitude spectrum is steeper than the original image, and also has a different shape. Figure D shows a sinc filtered version of the original image. The shape of the amplitude spectrum is radically different. This manifests itself as banding in the resulting image.

There are different techniques for introducing blur to natural images. consider Figure 4.8. The top row shows the images, and the lower row their amplitude spectra. Figure 4.8A the original natural image, and its amplitude
spectrum. Blur was introduced in Figure 4.8B by increasing the slope of the amplitude spectrum in frequency space. Artificially blurring natural images often involves filtering techniques such as Gaussian or sinc filtering. Figure 4.8C shows a Gaussian blurred images, and the corresponding amplitude spectrum. Gaussian filtering has been used in the past as a model of optical blur (e.g. Bocheva and Mitriani, 1993). However, this is limited as a model of optical blur, as it does not contain the correct phase relationships that result from optical phenomena such as aperture effects. Aperture effects result from optics. Light passes through the pupil, and is subjected to diffraction. This creates characteristic phase reversals in the resulting retinal image. These phase reversals appear as banding around the edges in the image. Sinc filtering, which also introduces these phase reversals, has been suggested to be a potentially more realistic model of optical blur than Gaussian blur (Murray and Bex, 2010). The difference between Gaussian and sinc filtering can be seen in Figure 4.8. As can be seen in Figure 4.8; the different filters affect the shape of the amplitude spectrum and the local contrast gradients. Whether this distinction matters or not depends on exactly what information the visual system uses to extract information from the image, or to drive ocular motor responses, such as the accommodation (focussing) response.

**Accommodation Responses**

Many images will naturally consist of relatively large amounts of low spatial frequency information, for example images with large sky or shadow regions (e.g. Tatler, Baddeley, and Gilchrist, 2005). Nevertheless, images with predominantly more low spatial frequency information tend to be perceived as blurred. Defocus blur is characterised by a loss of high spatial frequency information on the fixated object. This is a possible cue to drive the accommodation (focussing) responses -this blur can be reduced by focussing at the appropriate distance. It is thought that the goal of the accommodation response is to maximise retinal image contrast, but unfortunately, there is as yet a limited understanding of exactly what spatial frequency information the accommodative system uses (MacKenzie et al., 2010). There is also uncertainty about whether the system uses local or global image statistics to achieve this end. MacKenzie et al. (2010) have argued that the global amplitude of spatial frequency information is critical for driving accommodation responses: they consider only the Fourier amplitude spectrum in their model of how stimulus spatial frequency content affects accommodation responses. By contrast, Day et al. (2009a) make the case that the critical information for accommodation responses is contained in the steepness of the contrast gradients. Defocus
blur has been argued to be sufficient to drive accommodation responses in the absence of other cues (Horwood and Riddell, 2008). It has been suggested that coarse accommodation responses might be driven by the low spatial frequencies, which are then refined by using the higher ones (Charman, 1979). This is important as insufficient high spatial frequency information might provide inadequate information to drive accommodation, potentially leading to uncertainty in the response.

**Microfluctuations**

The accommodative system might use small shifts in accommodation amplitude, of around 0.25 dioptres (microfluctuations), as a tool to inform the response (Niwa and Tokoro, 1998). These microfluctuations introduce a small amount of defocus, affecting the contrast gradients of the image. This information can then be used as a directional cue to maximise the contrast gradients of the retinal image. The shallower the contrast gradients in the original image, the larger a shift in defocus is needed to see a difference in the retinal image. Thus greater microfluctuations might be indicative of increased uncertainty. Larger microfluctuations have been shown in the accommodation response for sine waves than for square-waveforms (Day et al., 2009a), which could suggest that lack of the high spatial frequency information (in the sine wave) leads to difficulties accommodating.

**Wavefront Aberrations**

The spatial frequency information contained in the retinal image is complicated by the imperfect optics of the eye, known as wavefront aberrations. It has been argued that these aberrations provide a directional cue to the accommodation response (Bühren and Collins, 2006). This is important as defocus is not a directional cue. For example, if an object is in focus, both increasing and decreasing accommodative distance means that the object becomes more blurred. For an out-of-focus object, it is uncertain whether to accommodate closer or further away to bring the object into focus based on microfluctuations alone. Aberrations provide information relating to the direction of optimal focus as well. These aberrations are highly idiosyncratic, so the effects they have will vary widely between individuals (Thibos, Applegate, Schwiegerling, Webb, and Members, 2002). Also there is individual variation on the extent to which these aberrations are used (Gambra, Sawides, Dorronsoro, and Marcos, 2009). Aberrations can be measured for an individual and expressed as a set of Zernike polynomials (Thibos et al., 2002), which describe the deviation from
4.1. INTRODUCTION

a diffraction-limited (perfect lens) system.

Informative stimuli for the accommodation response

In summary, an informative stimulus to drive accommodation responses would provide sufficient amounts of the relevant (higher) spatial frequency information. Images with steep contrast gradients (sharp edges) are composed of both low and high spatial frequencies. Natural images have a $1/f^1$ amplitude spectrum, meaning they contain a range of spatial frequencies. Therefore it might be the case that there is an ideal statistic of $1/f^1$, which would contain all the information necessary to drive accommodation responses, and most natural images are close enough to this ideal value to contain sufficient information.

Conversely, if an image does not contain much contrast at the high spatial frequencies, then adding blur will not affect the stimulus to a great extent. In other words the contrast gradients of a poor stimulus for accommodation might not change much with defocus blur. This will render the defocus cue to accommodation uninformative, which might increase uncertainty in the accommodative response. This uncertainty might then lead to inefficient, inaccurate accommodation responses. One possible response to this is that the accommodative system will increase microfluctuations, as greater microfluctuations would be needed to provide sufficient contrast differences. In this manner, a stimulus lacking in high spatial frequency information (such as filtered noise with a manipulated amplitude spectrum) might appear uncomfortable: the stimulus lacks the initial high spatial frequency information to refine the accommodation response, and therefore the accommodative system is unable to focus properly.

Current experiment

The aim of this study is to investigate the relationship between relative spatial frequency content of images, visual discomfort and perceived blur. This was investigated first using controllable, artificial stimuli to show the effects of luminance profile manipulations on discomfort judgements, to determine whether loss of spatial frequency information, or change in luminance profile, was the deciding factor for any possible discomfort judgements.

The experiment was then extended to more complicated natural images, using two different models of optical blur to assess the effect of phase reversals. Also, images were chosen with a range of initial amplitude spectra, to see if deviation
from the original, or an ideal $1/f^1$ statistic, is critical.

The second aim of the study was to create a theoretical account of the effect of uncomfortable stimuli on the accommodation response using an idealised model of how these stimuli would appear as a retinal image. The aim was to predict the cue that the accommodative system would receive from these different stimuli.

4.2 General Methods

Methods used were similar for the experimental sections of this study. Therefore general methods for all of the psychophysical experiments will first be described.

Apparatus

Stimuli were presented at a distance of one metre, on a 21 inch Sony Trinitron monitor with screen resolution of 1680 x 1050 and a vertical refresh rate of 60 Hz. Images were created and displayed using MATLAB 7.1 (The Mathworks Inc., 2005) and the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997; Kleiner et al., 2007). The luminance response of the monitor was measured and calibrated using a Minolta Luminance Meter, LS-110 photometer. The luminance of the mid-grey background was 38.5 cd/m$^2$, the luminance range was from 3 to 71 cd/m$^2$. Head movements were stabilised by use of a chin rest.

Stimuli

All stimuli were 840 x 840 pixel images. These were spatially vignetted with a circular window of radius of 200 pixels, 5.71°. Outside of this radius the edges of this window fell off with a Gaussian profile of $\sigma$ 0.93° (width at half height was 552 pixels) to eliminate hard edges. The size of the visible pattern subtended approximately 8.53°. The background was held at a constant mid-grey luminance level.

Observers

All participants in all experiments in this study were reimbursed five pounds per hour for their time. The entire study was approved by the University of St Andrews Teaching and Research Ethics committee.
4.2. GENERAL METHODS

Procedure
A two-interval-forced-choice (2IFC) task was used. Stimulus appeared on the screen for 0.6s each, with a delay of 0.1s between them, in which a mid-grey screen was presented. Stimuli were replaced with a mid-grey background during which the observer responded. Observers indicated which of the two stimuli appeared the more uncomfortable by pressing the corresponding left (first interval) or right (second interval) arrow key.

Analysis
Raw scores were relative discomfort judgements: the frequency a particular image or image category was chosen as more uncomfortable out of the paired comparisons. Raw scores as percentages from all experiments were normalised by taking the z-scores. This converts the raw results onto a scale of discomfort, which can then be directly compared.
4.3 Experiment 1: Discomfort Judgements of Simple Stimuli

Method

Stimuli:
Stimuli were circular gratings, consisting of a fundamental frequency of 0.375, 0.75 or 1.5 cycles/degree, plus up to four harmonics. There were two conditions, in which either the number of harmonics, or the relative phases of the harmonics were varied. Figure 4.2 shows examples of the stimuli and how they were created in terms of their luminance profile. In the first, harmonic manipulation condition, the number of harmonics was varied. In all cases, the fundamental frequency component was present. This was either presented alone, or accompanied by the third harmonic only; the third and fifth harmonics, the third, fifth and seventh harmonics or the third, fifth, seventh and ninth harmonics (see Figure 4.2). In all cases, the fundamental and all of the harmonics were in square-wave phase. In the second, phase manipulation condition, all the stimuli consisted of the fundamental frequency plus all four harmonics. The phase of the third harmonic was shifted relative to that of the fundamental by 0, 45, 90, 135, 180 degrees to create five different stimuli for each spatial frequency.

Observers:
14 naïve student observers took part in the experiment, all with normal or corrected to normal vision. Efforts were made to ensure that the same observers participated in both phase and harmonic conditions, however this was not always possible. Three out of the 14 observers participated in the harmonic condition only, and three out of the 14 in the phase condition only. Ages were between 18 and 29, mean age was 21.6 years.

Procedure:
Observers initially matched stimuli for perceived contrast using a self-adjustment technique. Stimuli were presented three times each in random order, and observers matched the perceived contrast to a standard. The standards were the sinusoidal stimulus (in the harmonic condition) and the 180° phase-shifted stimulus (in the phase condition). The Michelson contrast of these standards was 0.29; the contrasts of all other stimuli were perceptually matched to this. The Michelson contrast of the stimuli was equalised. However, RMS contrast was used as the metric by which to evenly extend the contrast of the image. Therefore, RMS contrast will be referred to later. The mean of these three
4.3. EXPERIMENT 1: DISCOMFORT JUDGEMENTS OF SIMPLE STIMULI

Figure 4.2: Stimulus luminance profiles

Top: example stimuli for the harmonic manipulated condition. Constituent harmonics are added together to create the total waveform. Total waveform is shown in bold. Resulting image is shown below. Lower row: example stimuli for the phase manipulated condition. Fundamental, third, fifth, seventh and ninth harmonics added together to create the total waveform, phase of the third harmonic is increasingly shifted relative to the other components. Resulting image is shown below.
4.3. EXPERIMENT 1: DISCOMFORT JUDGEMENTS OF SIMPLE STIMULI

Figure 4.3: Contrast matching results for Experiment 1

On the left is the average reduction in physical (RMS) contrast needed for the harmonic manipulated stimuli to be matched for perceived contrast. On the right is the average reduction in RMS contrast needed for phase manipulated images to be perceived as equal contrast. Error bars show $\pm 1$ standard error. Solid black line is the RMS contrast corresponding to the initial Michelson contrast of the standard.

settings was taken as the contrast setting for stimuli perceptually equal in contrast for that individual. Results of the contrast matching experiment can be seen in Figure 4.3. These are measurements of RMS contrast needed to be removed from the stimuli in order to match them for perceived contrast, averaged over all individuals.

Observers then took part in a two-interval-forced-choice (2IFC) task discomfort rating experiment. The contrasts of the stimuli were set to be perceptually equal, based on the responses of each observer in the previous part of the experiment. It should be noted that although average results are plotted in Figure 4.3, each individual was presented stimuli according to their own unique results from the contrast matching experiment. There were 20 repetitions of each comparison; each of the 15 stimuli was compared to all of the others.

Results

Contrast matching results in Figure 4.3 showed there to be variation in the contrast needed for each of the stimuli to be perceived as having equal contrast. In the harmonic manipulated condition more contrast was needed for stimuli
4.3. EXPERIMENT 1: DISCOMFORT JUDGEMENTS OF SIMPLE STIMULI

Left: Discomfort scores against increasing number of harmonics for three fundamental frequencies. There is a significant effect of harmonic manipulation on perceived discomfort; fewer harmonics are judged as more uncomfortable. Right: Discomfort scores against increasing phase difference of the third harmonic for three fundamental frequencies. There are no significant effects of either spatial frequency or phase difference.

with fewer harmonics compared to the square-wave profile stimulus. In the phase manipulated condition, the further out-of-phase a stimulus was, the less contrast is needed to make this appear equal in contrast to the others in the condition.

The left side of Figure 4.4 plots stimulus discomfort scores against increasing number of harmonics for each of the three fundamental frequencies. The results of a three (fundamental frequency) x five (number of harmonics) repeated-measures ANOVA showed a main effect of the number of harmonics ($F(4,52) = 6.91, p < 0.01$). There was no significant effect of spatial frequency ($F(2,26) = 0.22, p = 0.80$). There was no significant interaction ($F(8,104) = 1.33, p = 0.24$).

The right side of Figure 4.4 plots discomfort scores against increasing phase displacement of the third harmonic for three fundamental frequencies. The results of a three (fundamental frequency) x five (magnitude of phase shift) repeated-measures ANOVA showed there to be no significant effects of phase ($F(4,52) = 1.21, p = 0.32$), fundamental frequency ($F(2,26) = 1.01, p = 0.38$), or any interaction of fundamental frequency and phase ($F(8,104) = 0.84, p = 0.57$).
4.3. EXPERIMENT 1: DISCOMFORT JUDGEMENTS OF SIMPLE STIMULI

Discussion

The results showed an effect of the presence of high spatial frequency information on subjective discomfort judgements for simple stimuli. Removing high spatial frequency components (harmonics) increased discomfort. A concentration of visual information at low spatial frequencies has been associated with perceived blur in simple stimuli (Campbell, Howell, and Johnstone, 1978). This suggests that perceived blur and visual discomfort are related, either directly, in a causal relationship, or possibly that both are influenced by a common mechanism for both subjective attributes that depends on relative high spatial frequency information.

There was no significant effect of manipulations of the luminance gradient induced by increasing the phase difference of the third harmonic on perceived discomfort. This could not be explained by a difference in overall perceived contrast, as this was based on the results of the contrast matching task. Physical contrast also cannot account for the differences, as the stimuli with shallower contrast gradients (further out-of-phase) had less physical RMS contrast in order to match them.

The actual contrast gradients presented to each participant will have varied, depending on the contrast settings that were made. Figure 4.5 shows the average maximum contrast gradients of each of the stimuli presented. From Figure 4.5 it can be seen that there is no systematic difference between the phase and the harmonic condition stimuli. There is a reduction in maximum contrast gradient with decreasing number of harmonics and increasing phase shift, and there are differences between the fundamental frequencies, but overall there is no large difference between the two conditions. Therefore, the average presented contrast gradients also cannot account for the difference in discomfort judgements.

Shifting the phase introduces small variations in the overall contrast gradient known as ‘jaggies’. These are small deviations from a smooth contrast gradient due to the interactions between the different out-of-phase harmonics. These are thus small ‘notches’ in the overall gradient due to the presence of high spatial frequency information in the image. It could be argued that the presence of these ‘jaggies’ is the reason that these stimuli were not judged as uncomfortable. This explanation is consistent with the idea that lack of high

91
4.3. EXPERIMENT 1: DISCOMFORT JUDGEMENTS OF SIMPLE STIMULI

Average maximum contrast gradients of the presented stimuli. Stimulus contrast gradients varied over individuals: each was presented their own unique set of stimuli, matched for perceived contrast. H = harmonic manipulated condition, P = phase manipulated condition. 0.375, 0.75, 1.5 are the fundamental frequencies.
4.4. EXPERIMENT 2: BLUR JUDGEMENTS OF SIMPLE STIMULI

spatial frequency content is the important factor in visual discomfort.

The harmonic manipulations of this experiment affected discomfort judgements, and the phase manipulations did not. As both these manipulations affected the luminance gradients, it is expected that they may also affect perceived blur. However, this might be the case: luminance gradients can also be interpreted as shading (Kingdom, 2008). It is also unclear if the presence of ‘jaggies’ means that the phase-manipulated stimuli are not perceived as blurred, despite the shallower contrast gradient. Therefore in the second experiment observers were asked if the stimuli appeared more blurred, using the same setup as before.

4.4 Experiment 2: Blur Judgements of Simple Stimuli

The aim of Experiment 2 is to ascertain whether the stimuli identified as uncomfortable from Experiment 1 are actually perceived as blurred.

Method

Six undergraduate observers took part in the study. Two had taken part in Experiment 1, the other four were completely naïve to the purposes of the experiment. The contrast matching task was exactly the same as in Experiment 1. This time, participants were asked to choose the more blurred stimulus (not the more uncomfortable) of each pair, and indicate their response using the left and right arrow keys. They were asked to guess if they felt that neither stimulus was blurred.

Results

The left side of Figure 4.6 plots blur judgements against increasing number of harmonics. This shows there was an effect of the number of harmonics on perceived blur; the fewer harmonics were present, the more blur was perceived. The results of a three (fundamental frequency) x five (number of harmonics) repeated measures ANOVA showed there to be a significant effect of number of harmonics ($F(4,20) = 12.27, p < 0.01$). Stimuli with more harmonics are perceived as less blurred than stimuli with fewer harmonics. There was no significant effect of fundamental frequency ($F(1.8,9.0) = 2.28, p = 0.15$). There was a significant interaction between the number of harmonics and fundamental frequency on blur judgements ($F(8,40) = 4.65, p < 0.01$). Post-hoc one-way
4.4. EXPERIMENT 2: BLUR JUDGEMENTS OF SIMPLE STIMULI

repeated measures ANOVA showed there to be an effect of the number of harmonics for the lowest two fundamental frequencies ($F(4,20) = 24.28, p < 0.01$; and $F(4,20) = 7.50, p < 0.05$, for 0.375 and 0.75 respectively). However, there was no significant effect of number of harmonics for the highest fundamental frequency ($F(4,20) = 0.12, p = \text{NS}$).

The right side of Figure 4.6 displays the perceived blur against phase manipulation (degrees). There is no effect of the phase manipulation on perceived blur. The results of a three (fundamental frequency) x five (phase shift) repeated measures ANOVA showed there to be a significant effect of fundamental frequency ($F(2,10) = 9.12, p < 0.05$). There was no significant effect of phase shift ($F(4,20) = 0.22, p = 0.92$). There was no significant interaction between phase shift and fundamental frequency on blur judgements ($F(8,40) = 0.89, p = 0.53$).

![Figure 4.6: Blur judgements for Experiment 2](image)

Left: Blur judgements against number of harmonics for three fundamental frequencies. There is a significant effect of the number of harmonics: fewer harmonics increases perceived blur. There is also a significant interaction between the number of harmonics and fundamental frequency on perceived blur, meaning that the effect of the number of harmonics on perceived blur is greater for the lower fundamental frequencies of 0.375 and 0.75 cycles/degree.

Right: Blur judgements against phase difference of the third harmonic for three fundamental frequencies.

Discussion

There was a significant effect of the number of harmonics on perceived blur. As this is the same pattern of results as Experiment 1, this suggests that there
might be a relationship between perceived blur and discomfort judgements. There was also an interaction between number of harmonics and fundamental frequency on perceived blur judgements: the effect of harmonic removal was more evident for the lower fundamental frequency compared to the higher fundamental frequency. This could be because the contrast gradient of the highest spatial frequency fundamental is too steep by itself for the loss of higher spatial frequency information to have much effect.

The phase shift manipulation showed no effects on either discomfort judgements or blur judgements. This suggests that there is an effect exclusive to the harmonic condition. This could well indicate that the loss of high spatial frequency information results in both discomfort and blur. This is indicative that there might be a common mechanism underlying both these processes. There was an effect of fundamental frequency on perceived blur in the phase condition: stimuli with a higher fundamental frequency were seen as less blurred than stimuli with lower fundamental frequencies. This suggests that the initial steepness of the contrast gradient affects perceived blur.

4.5 Experiment 3: Blurring Natural Images

The first two experiments showed that, for simple radial square-wave stimuli, removing high frequency components increased both visual discomfort and perceived blur. Since the removal of high spatial frequency information also increases apparent blur in more complicated images (Webster et al., 2002; Murray and Bex, 2010), we next investigated whether this also leads to increased visual discomfort for complex natural images. Experiment 3 has three aims. The first is to investigate the effects of increasing blur on visual discomfort judgements using natural stimuli. The second is to investigate if there is a difference between phase manipulated sinc-filtered images and Gaussian-filtered images. The third is to see if visual discomfort judgements depend on deviation from the original statistics of the image, or deviation from a possibly ‘ideal’ $1/f^1$ statistic.

Method

60 greyscale natural images were taken from the Hibbard (2008) database. Images were 1201 x 1201 pixels in size, taken using a calibrated Nikon Coolpix 4500 digital camera. Calibrations corrected for lens distortion, focal length, and colour characteristics. Resolution was 1 pixel/arcmin. These were from two general categories: distant scenes (e.g. woodland scenes, beach scenes),
4.5. EXPERIMENT 3: BLURRING NATURAL IMAGES

or close-ups of vegetables, rocks and seaweed. 30 outdoor scene images had a mean exponent of value = -1.19 (standard deviation = 0.29). The 30 close-up images had a mean exponent of -1.10 (standard deviation = 0.20). The ten most extreme examples from each category were chosen for the test stimuli; the ten steepest slopes in the close-up category (mean = -1.39, SD = 0.06), the ten shallowest slopes in the outdoor scene category (mean = -0.95, SD = 0.06). These were significantly different according to an independent pairs t-test (t(18) = -28.39, p < 0.001).

Images were filtered in the Fourier domain with either a Gaussian or sinc filter. Gaussian filters were of varying sizes: $\sigma = 8, 16, \text{ or } 32 \text{ pixels in frequency space}$ (see Equation 1). This gave Gaussian filters of width at half height 20 pixels, 38 pixels, and 76 pixels, respectively.

$$A = k_g e^{(-\frac{f^2}{2\sigma^2})}$$  \hspace{1cm} (4.1)

Where $A$ is the amplitude of the filter, corresponding to the amount of the signal transmitted, $f$ is the frequency, $\sigma$ is the standard deviation in pixels, and $k_g$ is the normalisation constant.

Sinc filtered images were created by multiplying the Fourier-transform of the image with a sinc filter of varying size, resulting in three levels of blur (see Equation 4.2). $\lambda$ was chosen to be 2.9 times $\sigma$ values; 23, 46, 93. These levels of $\lambda$ were chosen as the perceived blur from these filters is comparable to the level of perceived blur from the Gaussian filtering (Murray and Bex, 2010).

$$A = k_s \frac{sin(\frac{zf}{\lambda})}{\frac{zf}{\lambda}}$$  \hspace{1cm} (4.2)

Where $\lambda$ is comparable to the $\sigma$ values of the Gaussian filter (see (Murray and Bex, 2010)). $k_s$ is the normalisation constant.

The two types of blurring (sinc and Gaussian) at three levels created six versions of each image. All images were matched for RMS contrast. The RMS contrast was set at 0.3 for all images in the current experiment. 13 naïve undergraduate students with normal or corrected-to-normal vision participated in the study. There were two conditions, one consisting of steep slope images,
the other of shallow slope images. 10 participants completed the steep slope condition first, three completed the shallow slope condition first. There were three levels of blur and two types of blur (sinc or Gaussian) for each of the ten pictures in each condition. This led to a total of six different image categories, with ten possible photographs of each. Each of the six categories was compared to each of the other categories, resulting in 15 possible comparisons. Each basic stimulus was presented only once, in all six photographs (one basic image, three levels of sinc filtering and three levels of Gaussian filtering). Ten images of each category were used, so that there were 10 test images × 10 standard images × 15 comparisons = a total of 1500 trials. The number of times a stimulus category was chosen as more uncomfortable was recorded.

Results

The left side of Figure 4.7 plots discomfort judgements against $\sigma$ (for the Gaussian filter; lower axis) or $\lambda$ (for the sinc filter; upper axis) values for the steep slope condition. There is a significant main effect of increasing blur (decreasing $\sigma$ or $\lambda$) (F(2,24) = 31.12, p < 0.01); increasing blur (decreasing $\sigma$ or $\lambda$) resulted in increased discomfort judgements. There was no significant effect of type of blur (Gaussian or sinc) (F(1,12) = 1.47, p = 0.25). There was no significant interaction between type of blur and amount of blur (F(1.7,21.1) = 0.61, p = 0.55).

The right side of Figure 4.7 plots discomfort judgements against $\sigma$ (lower axis) and $\lambda$ (upper axis) for the shallow slope images. There is a significant main effect of increasing blur only (F(2,24) = 12.61, p < 0.01); there were no effects of type of blur (F(1,12) = 0.73, p = 0.41) and no interaction (F(1.7,21.1) = 1.38, p = 0.27).

Discussion

There is a strong effect of blur on discomfort judgements in natural images - the more blur the more often the stimulus is judged as more uncomfortable of the pair of images presented. The results of this experiment are consistent with Experiment 1: blur is associated with discomfort judgements, this time for natural as well as simple stimuli.

It was not possible to gauge exactly how much the blurring functions increased the initial slope exponent as it is not possible to obtain a good linear fit of a curved function, therefore it cannot be shown whether or not these slopes
4.5. EXPERIMENT 3: BLURRING NATURAL IMAGES

Figure 4.7: Discomfort judgements for Experiment 3

Left: Discomfort judgements against decreasing blur for natural images for the steep initial slope condition. Blur is comparable amounts of Gaussian or sinc blur. \( \sigma \) and \( \lambda \) values relate to frequency, therefore small values indicate large amounts of blur. Right: Discomfort judgements against decreasing blur for natural images for the shallow initial slope condition. Blur is comparable amounts of Gaussian or sinc blur. \( \sigma \) and \( \lambda \) values relate to frequency, therefore small values indicate large amounts of blur.

are now closer or further from the ‘ideal’ \( 1/f \) statistic. This is demonstrated in Figure 4.8 - it would not be possible to fit a linear slope to the amplitude spectra of the Gaussian or sinc manipulated images.

One potential reason for discomfort from blurred images could be because blurring the image impoverishes the feedback for the accommodation response: as the original image is blurred, this does not provide the necessary signal for the accommodative response to focus. The signal to accommodation will be investigated in the next study, using an idealised model of defocus blur based on the modulation transfer function of the eye.
4.6. AN IDEALISED MODEL OF DEFOCUS BLUR

Figure 4.8: Amplitude spectra of natural images.

Top row shows different image manipulations, the bottom row shows the corresponding amplitude spectra. Luminance amplitude is plotted against log spatial frequency (cycles/image). Points on the graph are individual pixels from the Fourier amplitude spectrum. These indicate contrast amplitude at each spatial frequency, sampled at all orientations. Figure A shows the original image and its amplitude spectrum. Figure B shows an image in which the slope exponent ($\beta$) of the original image (1.45) has now been increased to 2. This steeper slope results in increased amplitude at the lower spatial frequencies. Figure C shows a Gaussian blurred version of the original image. Again, the amplitude spectrum is steeper than the original image, and also has a different shape. Figure D shows a sinc filtered version of the original image. The shape of the amplitude spectrum is radically different. This manifests itself as banding in the resulting image.

4.6 An Idealised Model of Defocus Blur

Introduction

The first two experiments showed that image blur affects discomfort judgements for both simple and natural stimuli. It could be that blurred stimuli are poor cues to the accommodation system, as they lack the higher spatial frequency information required to drive the response. In this section the cue to the accommodative response will be investigated by modelling how the contrast of the retinal image changes with defocus for comfortable and uncomfortable stimuli.
4.6. AN IDEALISED MODEL OF DEFOCUS BLUR

Modulation transfer function

The cue to accommodation could be modelled by comparing the reduction in contrast of the retinal image to the contrast available in the incident image (e.g. Bennett and Quigley, 2011). The function relating incident contrast level to the retinal contrast level is the modulation transfer function (MTF) of the eye (Levi, 1974). Microfluctuations of the eye introduce small amounts of defocus blur, which could be used to adjust the accommodative response (Day et al., 2009a). A large change in contrast with this small defocus blur would provide a more informative cue to the accommodative response than would be provided by a stimulus that does not change much with defocus. The current aim is to investigate the change in maximum contrast gradient with defocus blur of some images, to see if there is a substantial change in contrast gradient with the addition of defocus. It is predicted that contrast gradients that change rapidly with defocus will be typical of comfortable images, as these will provide an informative cue for the accommodation system. Conversely it is expected that images whose contrast gradients show limited variation with the addition of defocus blur will be uninformative cues and thus provide a source of discomfort.

Wavefront aberrations

Defocus blur is not the sole cue to accommodation. Factors in addition to defocus blur affect the accommodation response, including pupil size (which varies with light intensity), wavelength of light, and critically, individual differences in transmittance of the eye. The human visual system is not diffraction-limited: imperfections in human optics, known as wavefront aberrations, scatter the incident light, thus affecting the MTF. These aberrations are specific to an individual observer, and it is highly likely that these are used as a cue to accommodation (Plainis, Ginis, and Pallikaris, 2005). Using specialist equipment such as an autorefractor, it is possible to measure the aberrations of an individual’s eye (Thibos et al., 2002). Using these data it is possible to model an individual’s MTF (MacKenzie et al., 2010).

Using such measurements it is also possible to model the typical wavefront aberrations of human eyes, (e.g. Bennett and Quigley, 2011). Salmon and van de Pol (2006) compiled data from various studies on the aberrations of over 2,500 human eyes, reported in the form of Zernike polynomials, which will be used in the current study to model average human aberrations. Because different studies had measured aberrations at different pupil sizes, generally using drugs to induce cycloplegia, the measurements were extrapolated by the...
4.6. AN IDEALISED MODEL OF DEFOCUS BLUR

authors for a range of pupil sizes, in order to facilitate comparisons. This included the 3mm pupil size, which is the size that will be used in this model, as this is typical for photopic conditions (Groot and Gebhard, 1952).

Method

The model incorporated Zernike polynomials for the average eye of 3mm pupil, based on the data from Salmon and van de Pol (2006), see Table 1.

<table>
<thead>
<tr>
<th>Zernike mode</th>
<th>Zernike coefficient (microns)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Z(3,-3)</td>
<td>0.019 ± 0.016</td>
</tr>
<tr>
<td>Z(3,-1)</td>
<td>0.021 ± 0.018</td>
</tr>
<tr>
<td>Z(3,+1)</td>
<td>0.015 ± 0.013</td>
</tr>
<tr>
<td>Z(3,+3)</td>
<td>0.014 ± 0.013</td>
</tr>
<tr>
<td>Z(4,-4)</td>
<td>0.050 ± 0.004</td>
</tr>
<tr>
<td>Z(4,-2)</td>
<td>0.030 ± 0.003</td>
</tr>
<tr>
<td>Z(4,0)</td>
<td>0.010 ± 0.008</td>
</tr>
<tr>
<td>Z(4,+2)</td>
<td>0.050 ± 0.005</td>
</tr>
<tr>
<td>Z(4,+4)</td>
<td>0.050 ± 0.005</td>
</tr>
<tr>
<td>Z(5,-5)</td>
<td>0.001 ± 0.001</td>
</tr>
<tr>
<td>Z(5,-3)</td>
<td>0.001 ± 0.001</td>
</tr>
<tr>
<td>Z(5,-1)</td>
<td>0.001 ± 0.001</td>
</tr>
<tr>
<td>Z(5,+1)</td>
<td>0.000 ± 0.000</td>
</tr>
<tr>
<td>Z(5,+3)</td>
<td>0.000 ± 0.000</td>
</tr>
<tr>
<td>Z(5,+5)</td>
<td>0.000 ± 0.000</td>
</tr>
</tbody>
</table>

Table 4.1: Zernike modes and mean Zernike coefficients (with standard deviations) for a population of human eyes. See Salmon and van de Pol (2006)

The Zernike coefficients are the scaling factor to multiply the relevant Zernike polynomial to describe the deviation from a perfect lens. The RMS of this total deviation is the wavefront aberration, see Equation 4.3, taken from Thibos, Himebaugh, and Coe (2006).

\[ RMS = \sqrt{\frac{1}{N} \sum (C_n^m)^2} \]  

Where \( C \) is the Zernike coefficient, \( n \) and \( m \) are the indices for the Zernike modes. 200 ‘eyes’ were simulated by choosing Zernike modes from a normal distribution based on the above values. This was then used to create 200 simulated point-spread-functions (PSF). The PSF, taken from Goodman (1968)
was calculated, see Equation 4.4.

\[
PSF = \frac{1}{Ad^2\lambda^2} \left| FT\left\{ p(x,y) e^{-i\frac{2\pi}{\lambda} W(x,y)} \right\} \right|^2
\]  

(4.4)

Where \( p(x, y) \) is the simple equation for transmission through a pupil of diameter 3mm, \( \lambda \) is the wavelength of light (555nm), \( W(x, y) \) is the Zernike polynomials, and \( d \) is the focal distance of the eye (23mm, see Larsen (1979)). The use of only monochromatic light is a simplification. The model therefore does not include any possibility for assessing the effects of chromatic aberrations. Although there is some evidence for the use of chromatic aberrations to provide a directional cue to accommodation, this is not used by all untrained observers (Fincham, 1951). Additionally, there is still some debate surrounding whether or not all individuals use this information (see Kruger, Matthews, Aggarwala, and Sanchez, 1993), therefore it was considered justified to omit this from the present model. The amplitude of the Fourier transform of the PSF is the MTF. The MTF was then convolved with the image, by multiplying in the frequency domain. This process was repeated for each of the 200 simulated PSFs. Defocus was added by changing the Zernike coefficient of the relevant mode; the appropriate level of blur in dioptres is proportional to this coefficient (Thibos et al., 2006), see Equation 4.5.

\[
M = \frac{C^0_2 4\sqrt{3}}{R^2}
\]  

(4.5)

Where \( M \) is the level of defocus in dioptres, \( R \) is the pupil radius, and \( C^0_2 \) is the Zernike coefficient for the defocus mode.

The stimuli from Experiments 1 and 3 were used. The artificial stimuli were square-wave radial waveforms with a 0.375 cycles/degree fundamental. There were five of these waveforms; the fundamental alone, or with one, two, three or four harmonics, as in Experiment 1. The natural stimuli were created from one of the photographs from Experiment 2. As in Experiment 2, three levels of Gaussian blur were added: \( \sigma = 8, 16, \text{or } 32 \) pixels in frequency space. Both square-wave and Gaussian blurred natural images were normalised between 0 and 1. The stimuli were all vignetted with window of radius 200 pixels \( (3.33^\circ) \), with Gaussian slope edges of \( \sigma = 65 \) pixels \( (1.08^\circ) \).

Stimuli were convolved (multiplied in frequency space) with each of the 200
simulated MTF functions. The maximum contrast gradient of the image was calculated using the inbuilt MATLAB function ‘gradient’. This function numerically calculates the gradient from the value of one pixel to the adjacent one using linear interpolation, see Equations 4.6 and 4.7.

\[
\Delta I = \frac{\partial I}{\partial x} \hat{i} + \frac{\partial I}{\partial y} \hat{j}
\]  

(4.6)

Where \( I(x, y) \) is the intensity of the image and \( \hat{i} \) and \( \hat{j} \) are vectors in the \( x \) and \( y \) directions. The magnitude of the gradient was then calculated:

\[
|\Delta I| = \sqrt{\left(\frac{\partial I}{\partial x}\right)^2 + \left(\frac{\partial I}{\partial y}\right)^2}
\]  

(4.7)

The maximum value of this array was then found. The mean of the maximum gradient value for all 200 simulated image outputs was calculated for each level of defocus blur. Because the object of interest is the change in contrast gradients, the maximum contrast gradient of the defocused images was normalised using the focussed photograph of this image. This is not the aberration-free original image, but the focussed photograph in the sense of zero defocus blur, thus other higher-order aberrations are present. This process was repeated for all five harmonic manipulations and for the three levels of Gaussian blurred natural images. The maximum of \(|\Delta I|\) for each image was then found.

Additionally, the amplitude of the Fourier spectrum was also found for each of the 200 ‘eyes’, for each level of defocus. This was averaged over all 200 model ‘eyes’. Log amplitude was then taken as a more global measure of change in contrast with increasing defocus.

**Results**

The left side of Figure 4.9 plots the change in maximum normalised contrast gradient against defocus for the harmonic manipulated stimuli. There is a steeper fall in maximum contrast gradient for the circular waveforms with increasing number of harmonics. The right side of Figure 4.9 plots the normalised maximum contrast gradient against defocus for the three levels of Gaussian blurred natural images. The more Gaussian blur (smaller \( \sigma \) value),
4.6. AN IDEALISED MODEL OF DEFOCUS BLUR

the smaller the change in maximum contrast gradient with increasing defocus.

Figure 4.9: Modelled change in contrast gradients

Left: normalised maximum contrast gradient against increasing defocus blur in dioptres for radial waveforms. 0.375 cycles/degree waveform with varying number of harmonics. Right: normalised maximum contrast gradient against increasing defocus blur in dioptres for Gaussian blurred natural images. Contrast gradient was normalised against a focussed version of the image.

Figure 4.10 shows the change in maximum contrast gradient for phase manipulated simple waveforms. From Figure 4.10 there appears to be a similar decline in maximum contrast gradient with increasing defocus blur. However, in Figure 4.10 it can be seen that the phase manipulated radial waveforms show a steeper change in contrast gradient with increasing defocus blur than the harmonic manipulated stimuli. Despite some slight spreading of the phase-shifted stimuli, these are all comparable to the square-waveform, unlike the reduction in harmonics, which makes a dramatic difference to the maximum contrast gradients.

As previously discussed, contrast gradients are not the only measure of blur. Experiments 1 and 2 show that the loss of amplitude at the high spatial frequencies is important to discomfort judgements, it could be important to investigate the loss of global spatial frequency information also. Figure 4.11 shows the reduction in the amplitude spectrum with defocus for a natural image (top) and a Gaussian blurred version of the same image (lower). The Gaussian blur was $\sigma = 8$ pixels in the frequency domain. The blue colour
4.6. AN IDEALISED MODEL OF DEFOCUS BLUR

Figure 4.10: Maximum contrast gradient of phase manipulated stimuli

Normalised maximum contrast gradient against increasing defocus blur in dioptres for phase manipulated radial waveforms. 0.375 cycles/degree waveform with varying phase shift of third harmonic.
indicates less amplitude compared to the red. There is more amplitude at the midrange (around 2.5 cycles/degree) spatial frequencies (lower left of the plot) in the focussed image compared to the Gaussian blurred image. This relative lack of amplitude might cause uncertainty in the accommodative system as there is less information.

Figure 4.11: Natural image amplitude spectrum (log) change with defocus blur

Top: plot showing log of amplitude spectrum of natural image with increasing levels of defocus blur. Lower: plot showing log of amplitude spectrum of a Gaussian blurred ($\sigma = 8$) natural image with increasing levels of defocus blur.

The top image of Figure 4.12 shows the amplitude spectrum of a square-wave circular waveform with increasing level of defocus blur. The lower image of Figure 4.12 shows the amplitude spectrum of a sine wave stimulus with increasing levels of defocus blur. In a similar manner as the Gaussian blurred image, there is relatively less amplitude in the lower left of the sine-wave amplitude spectrum compared to the square-wave. Again, this is apparent at small levels of defocus up to around 5 cycles/degree spatial frequency - there is amplitude in the square-wave that is not present in the sine-wave. This indicates a possible lack in spatial frequency information in the two images. However, this time the difference is more subtle than with the natural images.
4.6. AN IDEALISED MODEL OF DEFOCUS BLUR

Figure 4.12: Sine wave radial waveform amplitude spectrum (log) change with defocus blur.

Plot showing log of amplitude spectrum of a sine wave radial waveform image with increasing levels of defocus blur.

The top figure in Figure 4.13 again shows the amplitude spectrum of square-wave (in phase image), the lower figure shows the amplitude spectrum of the 135 degree phase manipulated stimulus with increasing defocus blur. It can be seen from this figure that the amplitude spectra of these two are very similar, which is to be expected as the original images have the same spatial frequency content.

Figure 4.14 shows the amplitude change with defocus blur, averaged over all the spatial frequencies. From Figure 4.14 it can be seen that there is variation in the overall amplitudes between the different classes of image. The simple waveforms are all of higher overall amplitude than the natural images. The sine-wave has a flatter gradient than the square-wave at the small levels of defocus blur. Similarly the Gaussian image has a flatter gradient than the fo-
4.6. AN IDEALISED MODEL OF DEFOCUS BLUR

Figure 4.13: Phase manipulation amplitude spectrum (log) and defocus.

The change in log of amplitude spectrum of a five component complex waveform with the addition of defocus.

cussed image when small levels of defocus blur are added. This could indicate that amplitude of the low spatial frequencies, and how this changes with the introduction of small amounts of defocus blur, is potentially important to the accommodative response.

The overall amplitude could be important as the images were all normalised to begin with, and have the same amount of defocus blur added to them. The overall amplitude might therefore help distinguish between classes of stimuli, which is a point that will be addressed later in Chapter 9.

Discussion

Results of the model predictions to some extent support the idea that stimuli that are judged as more uncomfortable are less informative for the accommodative response. The simple stimuli judged as uncomfortable in Experiment 1 showed less of a change in maximum contrast gradient with defocus than those judged more uncomfortable. Additionally, a similar pattern was
4.6. AN IDEALISED MODEL OF DEFOCUS BLUR

Overall amplitude spectrum change with defocus averaged over all spatial frequencies.

found for natural images: those that were judged as more uncomfortable in Experiment 3 were also those showing less of a change in maximum contrast gradient with increasing defocus. This could be indicative that uncomfortable stimuli, which are perceived as blurred, do not provide an informative cue to the accommodation system, as there is a lack of change in the contrast of the stimulus with defocus.

The change in contrast was also apparent with the metric of the global amplitude spectrum. The global amplitude spectrum changes with the addition of defocus blur for any of the images tested: Gaussian blurred natural images, harmonic manipulated and phase manipulated stimuli. Phase manipulated stimuli and square-waveforms all had the same pattern of change with defocus in the amplitude spectrum. This is expected as they all had the same spatial frequency content.

What was interesting was the lack of midrange spatial frequency amplitude for the sine-waveforms and the Gaussian blurred images, compared to the square-wave and natural images respectively. This is in agreement with the lack of change in maximum contrast gradient from these stimuli. It is also indicative
that this information, which is thought to be important to the accommodative response (e.g. Charman, 1979), is lacking in these images.

4.7 Conclusion

Experimental Findings

The aim of this study was to investigate the relationship between low spatial frequency information, visual discomfort and perceived blur. Initial experiments investigated this relationship using simple stimuli. Experiment 1 showed that loss of high spatial frequency information, but not the shifting of the phase, resulted in increased discomfort judgements. A subsequent experiment showed that the same stimuli judged as more uncomfortable were also judged as more blurred, suggesting that there was an association between the loss of high spatial frequency information, discomfort and image blur. The relationship between loss of high spatial frequency information, discomfort and blur was further investigated using natural stimuli. Increasing blur led to increasing discomfort judgements, for both sinc and Gaussian blurred natural images, which supported the finding of the initial experiments using artificial stimuli. Phase manipulations also show this pattern, but did not affect discomfort judgements in the experimental part of this chapter.

Soso, Lettich, and Belgum (1980) conducted an experiment investigating paroxysmal activity in epilepsy sufferers on viewing striped patterns of different waveforms. Waveforms used were a basic sine wave of varying spatial frequency, as well as complex waveforms consisting of the initial sinewave and the third harmonic. The third harmonic was either in phase (square wave) or out of phase (triangular wave) with the fundamental. The sine wave elicited less activity than the complex waveforms, however there was no difference found in activity elicited by the square and triangular waveforms. The lack of difference between the activity elicited by the complex waveforms is in accordance with the current study. In the case of epilepsy, it is thought that excessive activity in the cortex might be the cause of difficulties from striped gratings (Wilkins et al., 1975). The finding of (Soso et al., 1980) that the sine wave elicits less epileptiform activity suggests that the discomfort reported in the present study might not have the same origin as discomfort from striped gratings. This does not contradict the notion that in this sample, with these blurred stimuli, the effect of discomfort might be due to accommodation responses, rather than excessive cortical responses.
Accommodative Cues

It was hypothesised that these stimuli could be uncomfortable as they provide uninformative cues for the accommodation system. This could be due to lack of change in contrast gradients with change in defocus. Two different metrics were used to measure the signal to the accommodative response: the change in maximum contrast gradient and the Fourier amplitude spectrum. Both of these metrics showed a similar result - stimuli judged as uncomfortable were those that did not change much with the increase of defocus blur. This could be seen particularly as a lack of change in the midrange spatial frequencies with the addition of defocus blur, as there was little amplitude present initially. Stimuli judged to be more uncomfortable in Experiments 1 and 3 showed shallower changes in contrast for both metrics on adding defocus blur. Therefore, these stimuli might be a less informative cue to accommodation responses, as small perturbations in accommodation amplitude (microfluctuations) will not have much of an effect on the kinds of information the accommodation system might use (whether contrast gradients or global spatial frequency content).

Motor and Sensory Problems

One possible reason why poor stimuli to accommodation might be uncomfortable is due to muscular fatigue in the motor response. The accommodation response is enacted using muscles that control the shape of the lens. It might be the case that increased amplitude microfluctuations require more effort to perform than smaller ones, and over an extended period the muscles controlling the shape of the lens might simply become fatigued. Alternatively, discomfort could be due to a sensory discrepancy; a discrepancy between expected and obtained image statistics. The response to the stimuli that lack the higher spatial frequency information can be described as open-loop; the retinal image remains blurred regardless of accommodative motor responses. If such expectations are violated, for example by open-loop stimuli to accommodation, which do not change with defocus, then this might be uncomfortable because of increased demand on neural resources required to interpret the image. Additionally, other information regarding the image, such as vergence, and the frame of the monitor, would indicate that the appropriate level of accommodation has been reached when the observer focuses at the screen distance, therefore the visual system might expect sharp edges. However, the stimuli remain blurred. This might be computationally costly for the visual system to process as it attempts to resolve this discrepancy.

The current study cannot differentiate between these two possibilities. This
could be achieved by objectively measuring the accommodation response using an autorefractor. If blurred stimuli really do present a poor stimulus to accommodation responses, then this might manifest itself in the accuracy of accommodation response (lead or lag), or the variation of the steady state response (microfluctuations). Stimuli that do not change much with defocus blur might be expected to result in more accommodative lead/lag, and more variable microfluctuations. This could also be investigated using behavioural methods. Pinhole viewing can have the effect of negating the accommodative microfluctuations in some individuals (Morrison, Siedel, Strang, and Gray, 2010). If the microfluctuations were reduced in this manner, any residual differences in discomfort judgements from blurred vs unblurred stimuli could not be accounted for in terms of microfluctuations. This is left for future work.

Alternative Explanations

Throughout this study, the aim has been to investigate the association between low spatial frequency information, blur and discomfort and to propose a mechanism for this in terms of accommodative responses. It has been suggested that stimuli deviating from natural image statistics might not be informative cues to drive the accommodative response. The idea that deviations from natural image statistics relate to visual discomfort has already been proposed (Juricevic et al., 2010). These authors suggested that deviations from the statistics of natural images are uncomfortable due to inefficient neural coding in the brain, instead of ocular motor responses. One way in which this could be achieved is to try to ensure a sparse response to typical natural stimuli. A sparse encoding of images is one that produces a strong response in only a relative small proportion of neurons, and has been shown to be a useful characterisation of the coding of natural images in the primary visual cortex (e.g. Field, 1994; Olshausen and Field, 2004; van der Schaaf and van Hateren, 1996; Field, 1999). Stimuli resulting in inefficient (non-sparse) neural coding might be uncomfortable as they are again costly in terms of metabolic resources.

This is not to suggest that the accommodation and neural accounts are competing. There are many different aspects of visual discomfort, and these may well account for two separate aspects. For example, the blurred vision and eye-strain reported under some circumstances, such as prolonged reading (Sheedy et al., 2003), might be indicative of poor accommodative responses. By contrast, cortical explanations can better account for why other stimuli, such as sharp, high contrast square-wave patterns are uncomfortable to non-clinical populations (Wilkins et al., 1984), and can additionally elicit neural activity
4.7. CONCLUSION

typical of seizures in epilepsy sufferers (Wilkins et al., 1979). Future study would be directed towards discriminating between these two possibilities, potentially assessing the separable effects on clinical populations, and optics in non-clinical populations.
CHAPTER 5

DEPTH-OF-FIELD AND VISUAL DISCOMFORT

5.1 Introduction

This work was in collaboration with Dr. H. T. Nefs, and Miss T. Zhang, of the TU Delft.

The results of Chapter 4 showed that spatial frequency manipulations that increase perceived blur also increase discomfort judgements for filtered noise stimuli. These manipulations typically increase the relative amount of contrast amplitude at low spatial frequencies compared to that typical of natural images. However, the manipulations used in Chapter 4, such as filtering in the Fourier domain, decrease the amount of high frequency information at all locations in the image, and therefore do not allow the location of the blurred regions to be taken into account.

The location of the blur is important. Incorrectly focussed images are blurred at the point of fixation. For the purposes of clarity, this is termed ‘defocus’ here. Chapter 4 demonstrated that global image blur (i.e., at the point of fixation as well) could have implications for the accommodative response, and this could potentially account for discomfort. However a certain amount of optical blur is present even in correctly focussed natural images, and this depends on the depths of the other objects relative to the point of fixation. Objects at depths different from the focal point have varying degrees of sharpness, depending on how far away they are. The distance (in depth) away from the
5.1. INTRODUCTION

focal point of the image that is in acceptable focus is the depth-of-field (DOF).

Depth-of-Field Blur

Images typically contain objects at different depths, but the range of objects that are in focus occupies a band of depths specified by optics. Thus blur arises on objects at depths other than the fixated object (focal point). The width in depth of this band of focus depends on both the distance from the observer to the focal point (egocentric distance), and the aperture size (see Equation 5.1). \( b \) is the blur circle, an area of the image in the frontoparallel plane that is determined to be blurred. The width of this blur circle increases with increasing depth from the point of focus.

\[
b = A \frac{s_0}{d_0} \left| 1 - \frac{d_0}{d_1} \right| \tag{5.1}
\]

Where \( A \) is the pupil diameter, \( s_0 \) is the axial length, \( d_0 \) is the distance from the observer to the focal point, and \( d_1 \) is the distance from the observer to the blurred object (flanker) see (Held, Cooper, O’Brien, and Banks, 2010).

The effects of DOF blur are demonstrated in Figure 5.1. Consider non-fixated objects at different depth planes from the focal point. These become gradually more blurred with increasing depth from the plane of the focal point. DOF also depends in part on aperture size. A shallow DOF creates a large blurring effect with just a small depth difference between the focal point and non-fixated object, and results from close viewing distances and a large aperture (see Figure 5.1a). A large DOF is found on viewing through small apertures, which leads to all objects in an image being sharp (see Figure 5.1b).

Depth-of-Field as a Cue to Distance

As DOF depends on the distance from the observer to the focal point (termed here ‘egocentric distance’), DOF blur could conceivably be used as a cue to distance, as well as depth \(^1\). As previously stated, natural images typically

\(^1\)To note the terminology, there is a convention that ‘distance’ is from the observer to the focal point, and ‘depth’ is the distance between any two points in the scene. This will be upheld here.
5.1. INTRODUCTION

Figure 5.1: Shallow and large depth-of-field

Left: shallow depth of field, with large blur gradients. Although the background of the photograph on the left is mostly blurred, it is also properly in focus for the focal point. Right: large depth of field: much of the image is in focus, even the flanking objects, which are further away than the focal point. Photographs courtesy of Dr H. T. Nefs, (Nefs, 2011)

have a $1/f^\beta$ amplitude spectrum, however, $\beta$ shows slight variation around 1, ranging from 0.8 to 1.5 (Tolhurst et al., 1992). Intrinsic blur away from the focal point from DOF could be one reason why close-up images typically have steeper amplitude spectra compared to more distant scenes (Redies et al., 2007). DOF has more of an effect at closer viewing distances (see Equation 5.1). If observers are able to detect the amount of DOF blur that should be present at a given viewing distance, they might expect more DOF blur for close-up images than more distant images.

Due to the variation of DOF blur with distance to the focal point, DOF could potentially provide a cue to egocentric distance: Vishwanath and Blaser (2010) showed that DOF blur affects estimates of egocentric distance. It has also been demonstrated that DOF blur can facilitate performance on psychophysical tasks using stereoscopic images, compared to completely sharp images: Banks, Akeley, Hoffman, and Girshick (2008) showed that observers were better at detecting the orientation of lines close to vertical when the distracting lines at other depths were blurred from DOF effects. Additionally, DOF can reduce rivalry from monocular regions in stereoscopic images (Hoffman and Banks, 2010). This indicates that DOF blur can not only be detected, but also used to influence performance when using stereoscopic displays, possibly by providing a distance and a depth cue (Held et al., 2010).

Datta, Joshi, Li, and Wang (2006) showed that aesthetic judgements of photographs are influenced by factors including DOF blur. The authors ascribed
5.1. INTRODUCTION

this to the required skill involved in creating these photographs. There is evidence that increased DOF increases both aesthetic quality and facilitates performance: Hillaire, Lécuyer, Cozot, and Casiez (2008) showed that adding blur gradients to simulate DOF cues facilitated performance on computer game playing for inexperienced players. Inexperienced players also judged the DOF blur gradients to add realism to the scenes. However, the performance of experienced players was not increased with the addition of these blur gradients. Additionally, some players judged the additional blur to increase fatigue and discomfort. Therefore, DOF can be judged as more aesthetically pleasing and facilitate performance, but can also hinder performance and impact discomfort judgements. In this study the effects of DOF blur gradients on discomfort are assessed.

Discomfort from Depth-of-Field

DOF blur itself does not drive the accommodation response (Day, Seidel, Gray, and Strang, 2009b). If it is the case that DOF manipulations alone affect discomfort, then the cause must be associated with blur itself, rather than accommodation responses. Therefore investigating the impact of DOF manipulations on discomfort judgements would allow discrimination between discomfort from blur or from accommodation responses. It is possible that the visual system can detect appropriate amounts of DOF blur for the depths in the scene given pupil size and an estimate of distance to the focal point. If so, then it is possible that inappropriate DOF blur gradients for a given pupil size and distance are detected by the visual system and cause discomfort. Additionally, as DOF is a cue to depth, it is possible that incongruent DOF blur might provide a source of conflict with other cues to depth, such as vergence and accommodation. If so, it is expected that the effects would depend on the distance to the focal point, as narrower DOF is expected at closer distances.

Additionally, DOF might be expected to influence discomfort resulting from an existing accommodation-vergence conflict. Vergence, defocus and DOF all provide cues to distance. The conflict between defocus blur and vergence when they specify different distances and depths has been shown to be a source of discomfort (Hoffman et al., 2008). It might be the case that if DOF provides an additional source of egocentric distance information, then this might affect the conflict, by strengthening one of the cues. Additionally, DOF blur would have the effect of reducing the disparity signal from the objects in the background, as detecting small disparities relies on high spatial frequency information (see Howard and Rogers (2002), pages 177-181 for a review of small disparity detec-
tion). Reducing the strength of the disparity cue by blurring the high spatial frequencies needed to provide disparity information might affect the strength of the accommodation-vergence conflict. This might therefore be expected to ameliorate discomfort from accommodation-vergence discrepancies.

**Current Study**

The aim of this study is to investigate the effect of DOF blur on discomfort judgements of stereoscopically presented images, with or without an additional accommodation-vergence cue conflict. In Condition 1, distance to the focal point will be defined by coherent accommodation and vergence cues: the focal point will remain at zero disparity, at the screen distance. DOF blur on the flanking objects will be varied, by changing the aperture on the camera used to take the pictures. This will result in DOF levels appropriate for human optics, or DOF that creates too much or too little blur in the image. It is predicted that the appropriate level of DOF for the assumed human optics will result in less discomfort than either too much or too little blur.

Condition 2 assesses a potential effect of DOF blur on pre-existing accommodation-vergence conflict. Accommodative distance will remain constant at the screen distance. Vergence-defined distance will vary with the depth of the objects in the scene, so that the focal point is projected to a closer distance than that specified by accommodation cues. In this condition the focal point will always be the one in sharp focus, however it will no longer be at zero disparity, thus creating the vergence-accommodation conflict typical of stereographic images. The aim of this condition is to ascertain whether accommodation-vergence induced discomfort is influenced by the level of DOF blur.

The effects of simple DOF and accommodation-vergence conflicts will also be directly compared using a within-subjects design, as all participants will complete both conditions.

**5.2 Method**

**Apparatus:** Images were photographed using a Nikon D90 digital camera with a fixed focal length lens of 50mm. Stimuli were presented on a Wheatstone stereoscope consisting of two 17inch NEC AccuSync LCD 72vm monitors, which were gamma corrected for colour using the Spyder pro 3 (Datacolor AG). Screen size was 34 x 27.5cm, and resolution 1280 x 1024 pixels. The distance to the screens was 53cm, viewed through two front-silvered mirrors.
5.2. METHOD

Figure 5.2: Stereoscope setup

The orientation of the mirrors was set so that the distance specified by vergence to the screen was the same as the accommodation-defined distance (zero disparity). Images were presented and displayed using MATLAB 2006 (The Mathworks Ltd) and the Psychtoolbox (Brainard, 1997; Pelli, 1997; Kleiner et al., 2007). A diagram of the stereoscopic experimental setup is seen in Figure 5.3. Head movements were stabilised using a chinrest.

Stimuli:
Photographs of objects were taken in a lightbox to make the colour stereo half images. Illumination was from overhead D65 840 lux daylight bulbs using a lightbox (Verivide Ltd, Leicester). Central objects could be either: a miniature book, a wind-up toy, or a packet of playing cards. Flanking objects were either: packets of tissues, highlighter pens, or stationary packets (paper-clips and drawing-pins). Layout of objects can be seen in Figure 5.2. Colour stereo half images in JPEG format were resized to be 550 x 880 pixels using bilinear interpolation, which resulted in an image 14.5cm high, thus subtending 15.3° of visual angle vertically. Photographs were spatially vignetted with a circular window of radius 300 pixels. Outside of this radius the edges of this window fell off with a Gaussian profile of $\sigma = 1.81^\circ$ to eliminate hard edges. The mean luminance of the images was 3.8cd/m². The central fixation point was presented at zero disparity (the screen distance). Two flankers appeared either side of the fixation point. The distance between the two flankers was always 5°, thus they moved further apart with increasing distance. These had a dispar-
5.2. **METHOD**

ity defined depth from the central fixation point (flanker-fixation separation). This corresponded to distances of 60cm, 70cm or 80cm back from the observer. This was created by the disparity from the two half-image photographs.

There are three levels of DOF blur for each flanker distance; too little, appropriate, or too much for the corresponding disparity-defined distance. How much this additional blur would correspond to in depth can be seen in Table 5.1. This was calculated using Equation 5.1. These calculations were based on an assumed pupil size of 4.5mm, which is expected given the presentation luminance of approx. 3cd/m\(^2\) average (Groot and Gebhard, 1952). Images were made by taking photographs of the setup using the corresponding f-stop on the camera to add the amount of blur for that stimulus; f22, f11, and f4. Each block contains nine possible scenes, and three possible flanker distances (60, 70, 80cm). This results in nine scenes x three flanker distances x three repetitions = 81 presentations per block. The order of completing the three blocks was counterbalanced. Stimulus f-stop, blur circle, corresponding flanker distances, and the appropriate theoretical distance for that particular amount of DOF blur, are shown in Table 5.1.

From Table 5.1 it can be seen that the blur circle from the optics of the eye (second column) is similar to the f-stop f11 (fourth column) with a camera of fixed focal length of 50mm. Too much blur is added by taking the photograph with the same object set-up, but a different aperture (f4). Too little blur is added by taking the photograph with the smaller f-stop (f22). For this amount of blur, given the optics of the eye, this would correspond to the final column of flanker distances. For example, with too much blur (resulting from a larger aperture than the eye, such as f4) the DOF would define a flanker distance of 78.25cm, whereas the disparity-defined flanker distance remains at 60cm. From this a cue conflict is created between disparity-defined distance (60cm) and DOF defined distance (78.25cm).

In Condition 1 stimuli were photographed with consistent accommodation-vergence cues. In Condition 2 stimuli with conflicting accommodation-vergence cues were created by taking photographs of the objects with the cameras focused at the distance of the central fixation point, but converged at the centre of the two flankers, not at the fixation point. This gave the central fixation
Figure 5.3: Object layout
5.2. METHOD

<table>
<thead>
<tr>
<th>Flanker distance (cm)</th>
<th>Eye (theoretical) blur circle (x10^-4 °)</th>
<th>f-stop</th>
<th>Camera (actual) blur circle (x10^-4 °)</th>
<th>Corresponding flanker distance (cm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>60</td>
<td>0.06</td>
<td>4</td>
<td>0.16</td>
<td>78.25</td>
</tr>
<tr>
<td>60</td>
<td>0.06</td>
<td>11</td>
<td>0.06</td>
<td>60.02</td>
</tr>
<tr>
<td>60</td>
<td>0.06</td>
<td>22</td>
<td>0.03</td>
<td>56.30</td>
</tr>
<tr>
<td>70</td>
<td>0.12</td>
<td>4</td>
<td>0.33</td>
<td>163.00</td>
</tr>
<tr>
<td>70</td>
<td>0.12</td>
<td>11</td>
<td>0.12</td>
<td>69.90</td>
</tr>
<tr>
<td>70</td>
<td>0.12</td>
<td>22</td>
<td>0.06</td>
<td>60.04</td>
</tr>
<tr>
<td>80</td>
<td>0.16</td>
<td>4</td>
<td>0.46</td>
<td>800.00</td>
</tr>
<tr>
<td>80</td>
<td>0.16</td>
<td>11</td>
<td>0.17</td>
<td>80.40</td>
</tr>
<tr>
<td>80</td>
<td>0.16</td>
<td>22</td>
<td>0.08</td>
<td>60.39</td>
</tr>
</tbody>
</table>

Table 5.1: Table of DOF blur calculations.

The first column is the distance from the observer to the flankers. The second column shows how big the blur circle would be for this distance, given the optics of the eye (given a human 4.5mm pupil), when the observer is focused on the fixation point. The fixation point is at 53cm distance from the observer. The third column lists the aperture sizes (f-stops) that were used to make the differing levels of DOF blur. The level of blur in the resulting photograph from this aperture and distance combination is given in the fourth column. The corresponding distances that the flankers would have to be at in order to result in this level of blur with human optics is in the final column.

point a disparity, such that it moved forwards in depth from the screen. The flankers would therefore be presented at zero disparity (screen distance), but be blurred as a result of DOF blur gradients. Despite the blur, the flankers do not provide a cue to the accommodation system as the observer is required to fixate the central object, not the flankers. Blur from DOF does not affect the accommodation response directly - a drop in overall luminance on the central target as a result of decreasing DOF affects the accommodation response, but only below a certain luminance level (around 0.1 cd/m², much lower than in the current experiment) (Day et al., 2009b).

Red square dots were added to the central fixation object. A random number of dots between 9 and 21 were generated. These were 10 pixels, corresponding to 0.28° in width. They spanned a radius of 150 pixels, or 4 degrees across the central focal object. In both conditions, these were given a disparity to match the central fixation object. In condition 1 these were all at zero disparity. In condition 2, dots were given disparity to match the central fixation point, thus they appeared to be at the depth of the target. Due to the object curvature, this could not be set to be the same for all images. The theoretical disparity of each fixation object depended on the depth of the flankers: 0.40°, 0.84°, and
1.16° disparity to correspond to 7cm, 17cm and 27cm depth. Theoretically the dots should have been the same, apart from the curvature of the objects meant there needed to be slight variation, otherwise the dots appeared to float within the object. This was done by measuring the disparity between the left and right half images in pixels, and moving the dots to appear on the central focal object. This was checked individually for each image, as there were slight variation between the images: overall mean disparity for the central targets was 0.97° (SD=0.74°). Again, there were nine scenes of three disparity defined fixation point differences, repeated three times within each block. There were again three blocks for each f-stop.

Observers: 22 corrected-to-normal naïve young observers (age = 22.7 years, SD = 3.7), participated in the study. All observers were screened for stereoacuity using the TNO test (Lameris Ootech B.V.).

Procedure: A white fixation cross against a mid-grey background preceded each trial. Trials were initiated by pressing the ‘enter’ key. Stimuli were presented for 5 seconds each. The fixation cross was presented on the mid-grey background between each stimulus presentation, which remained on screen until the observer made their response using the number keys. The observers were to count the dots on the central focal point (dots added later to the photographs), the number of which varied randomly on every trial, between 9 and 21. This was to encourage observers to fixate the focal object, and also to facilitate concentration. Each f-stop was presented in one block. Observers were presented with the three repetitions of the nine different scenes at three different distances in random order within a block. Observers completed the blocks in different orders to counterbalance any practice effects.

At the end of each block each observer completed a short questionnaire based on the work of Hoffman et al. (2008). Observers rated discomfort, dry eyes, headache, eyestrain, distortion and clarity of vision on a 10-point Likert scale. The dry eyes question was not included in the original Hoffman et al. (2008) questionnaire. This was added as there is research to indicate that this is a separate cause of discomfort (Sheedy et al., 2003). Therefore this is included as a control condition, and is not expected to vary with the experimental manipulations. The question on distortion was also not originally from the Hoffman et al. (2008) questionnaire. This was included as observers might have noticed that the level of blur was incorrect for the depths defined by disparity, and this question was to address this. The discomfort question was calibrated by asking observers to compare a score of 10 to staring at a bright
5.2. METHOD

light, for example see Table 5.2. Additional explanation was given to the
distortion question also. This was to give the participants some indication
of what was meant by these terms, and also to allow them to gauge their
judgements against something tangible. At the end of each condition there
was a questionnaire of free responses relating to all three blocks.

<table>
<thead>
<tr>
<th>Discomfort Questionnaire</th>
</tr>
</thead>
<tbody>
<tr>
<td>Please rate the images you just saw for:</td>
</tr>
<tr>
<td>Discomfort</td>
</tr>
<tr>
<td>1=no discomfort, 10=aversive, like staring at a bright light</td>
</tr>
<tr>
<td>1 2 3 4 5 6 7 8 9 10</td>
</tr>
<tr>
<td>Do your eyes feel dry?</td>
</tr>
<tr>
<td>1=no dryness, 10=very dry</td>
</tr>
<tr>
<td>1 2 3 4 5 6 7 8 9 10</td>
</tr>
<tr>
<td>How does your head feel?</td>
</tr>
<tr>
<td>1=no headache, 10=headache</td>
</tr>
<tr>
<td>1 2 3 4 5 6 7 8 9 10</td>
</tr>
<tr>
<td>How do your eyes feel?</td>
</tr>
<tr>
<td>1=no eyestrain, 10=eyes ache lots</td>
</tr>
<tr>
<td>1 2 3 4 5 6 7 8 9 10</td>
</tr>
<tr>
<td>Do you experience distortion in the images?</td>
</tr>
<tr>
<td>1=no distortion, 10=like looking in a funfair mirror</td>
</tr>
<tr>
<td>1 2 3 4 5 6 7 8 9 10</td>
</tr>
<tr>
<td>How clear is your vision?</td>
</tr>
<tr>
<td>1=very clear, 10=very blurry</td>
</tr>
<tr>
<td>1 2 3 4 5 6 7 8 9 10</td>
</tr>
</tbody>
</table>

Table 5.2

After all three experimental blocks, observers complete a final questionnaire
(see below). This is based on that of (Hoffman et al., 2008).

**Final Questions**
Which session was most fatiguing?
Which session irritated your eyes the most?
If you felt headache, which session was the worst?
Which session did you prefer?


5.3 Results

Analysis

Four observers’ data had to be excluded from analysis due to failing the TNO test. One observer’s dot counting data were lost due to a technical error during the experiment, although the participant still completed the written discomfort ratings. Initially, all data regarding subjective ratings were analysed together. A second analysis was performed which removed from the analysis those who reported no discomfort. No discomfort was defined as those who did not report any score higher than three on the ten point scale. This resulted in 13 observers to be included in the analysis. Mean rating of these 13 overall judgements was 3.56, SD = 2.00. These results were normalised by converting individual scores into z-scores for that individual. This was to reduce the effect due to individual variation in likelihood of responding positively. Results from the dot counting task were analysed by calculating the unsigned difference between the number of estimated dots and the actual number of dots. The average of this was taken to obtain a measure of accuracy at the task.

Subjective Ratings

Raw Scores

Discomfort judgements

Figure 5.4A shows discomfort judgements against decreasing aperture (increasing DOF, decreasing blur) for photographs with DOF-vergence only (Condition 1) and DOF-accommodation-vergence (Condition 2) conflict. A 2 (with or without accommodation-vergence conflict) x 3 (aperture) repeated measures ANOVA showed there to be no significant effects (conflict level F(1,17) = 1.66, p = NS; aperture F(1,31.5) = 2.41, p = NS; conflict-aperture interaction F(2.0,33.8) = 0.14, p = NS).

Dry eyes

Figure 5.4B shows ratings of dry eyes against increasing aperture (decreasing DOF) for DOF-vergence only and DOF-accommodation-vergence conflict. A 2 x 3 repeated measures ANOVA showed there to be no significant effects (conflict level F(1,17) = 1.19, p = NS; aperture F(1,33.1) = 0.67, p = NS; conflict-aperture interaction F(1.5,25.3) = 0.20, p = NS).

Headache

Figure 5.4C shows headache ratings against decreasing aperture size (increas-
5.3. RESULTS

Figure 5.4: Discomfort judgements against decreasing aperture.

Increasing DOF, decreasing image blur with increasing aperture for each of the ratings. Error bars show ±1 standard error.

ing DOF and decreasing blur). Results of a 2 x 3 repeated measures ANOVA showed no significant effects of either conflict level (F(1,17) = 0.50, p = NS); aperture (F(1.6,26.7) = 0.34, p = NS); or a conflict-aperture interaction (F(1.7,28.1) = 0.47, p = NS).

Eyestrain Figure 5.4D shows eyestrain ratings against decreasing aperture size (increasing DOF and decreasing blur). Results of a 2 x 3 repeated measures ANOVA showed no significant effects of either conflict level (F(1,17) = 1.36, p = NS); aperture (F(2,34) = 0.02, p = NS); or a conflict-aperture interaction (F(2,34) = 1.46, p = NS).

Distortion

Figure 5.4E shows distortion ratings against decreasing aperture size (increasing DOF and decreasing blur). Results of a 2 x 3 repeated measures ANOVA showed no significant effects of either conflict level (F(1,17) = 0.23, p = NS); aperture (F(1.8,31.3) = 0.15, p = NS); or of a conflict-aperture interaction (F(1.9,33.0) = 0.49, p = NS).
5.3. RESULTS

<table>
<thead>
<tr>
<th>F-stop</th>
<th>fatiguing</th>
<th>irritated</th>
<th>headache</th>
<th>prefer</th>
</tr>
</thead>
<tbody>
<tr>
<td>f4</td>
<td>8</td>
<td>7</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>f11</td>
<td>9</td>
<td>7</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>f22</td>
<td>2</td>
<td>5</td>
<td>7</td>
<td>8</td>
</tr>
</tbody>
</table>

Table 5.3: Condition 1 table of frequencies

<table>
<thead>
<tr>
<th>F-stop</th>
<th>fatiguing</th>
<th>irritated</th>
<th>headache</th>
<th>prefer</th>
</tr>
</thead>
<tbody>
<tr>
<td>f4</td>
<td>5</td>
<td>6</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>f11</td>
<td>11</td>
<td>9</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>f22</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>8</td>
</tr>
</tbody>
</table>

Table 5.4: Condition 2 table of frequencies

Clarity
Figure 5.4F shows clarity judgements (increasing blur) on the y axis against decreasing aperture size (increasing DOF and decreasing blur) on the x axis. A 2 x 3 repeated measures ANOVA showed no significant effects of either conflict level (F(1,17) = 1.86, p = NS); aperture (F(1.7,28.2) = 0.62, p = NS); or of a conflict-aperture interaction (F(2.0,33.6) = 1.82, p = NS).

Final Questions
The results of the final questions at the end of each block were also analysed. The frequency of observers who chose each block (f-stop) as answering a particular question was recorded. These frequencies can be seen in Table 5.3 for Condition 1 and Table 5.4 for Condition 2.

Results of a $\chi^2$ test showed there to be no significant result from the final four questions. No particular aperture appeared in any of the four categories more than by chance for either the DOF only condition ($\chi^2(6) = 7.90$, p = NS) or for the accommodation-vergence condition ($\chi^2(6) = 7.02$, p = NS).

Normalised Scores
Discomfort judgements
Figure 5.5A shows discomfort judgements against decreasing aperture (increas-
5.3. RESULTS

ing DOF, decreasing blur) for photographs with DOF-vergence only (Condition 1) and DOF-accommodation-vergence (Condition 2) conflict. A 2 (with or without accommodation-vergence conflict) x 3 (aperture) repeated measures ANOVA showed there to be no significant effects (conflict level $F(1,12) = 1.86$, $p = \text{NS}$; aperture $F(1.7,21.3) = 1.83$, $p = \text{NS}$; conflict-aperture interaction $F(2,24) = 0.33$, $p = \text{NS}$).

**Dry eyes**

Figure 5.5B shows ratings of dry eyes against increasing aperture (decreasing DOF) for DOF-vergence only and DOF-accommodation-vergence conflict. A 2 x 3 repeated measures ANOVA showed there to be no significant effects (conflict level $F(1,12) = 1.25$, $p = \text{NS}$; aperture $F(1.9,23.1) = 1.19$, $p = \text{NS}$; conflict-aperture interaction $F(2,24) = 0.01$, $p = \text{NS}$).

**Headache**

Figure 5.5C shows headache ratings against decreasing aperture size (increasing DOF and decreasing blur). Results of a 2 x 3 repeated measures ANOVA showed no significant effects of either conflict level ($F(1,12) = 0.62$, $p = \text{NS}$); aperture ($F(1.5,18.0) = 0.61$, $p = \text{NS}$); or a conflict-aperture interaction ($F(1.5,18.4) = 0.25$, $p = \text{NS}$).

**Eyestrain**

Figure 5.5D shows eyestrain ratings against decreasing aperture size (increasing DOF and decreasing blur). Results of a 2 x 3 repeated measures ANOVA showed no significant effects of either conflict level ($F(1,12) = 2.02$, $p = \text{NS}$); aperture ($F(1.4,17.2) = 0.02$, $p = \text{NS}$); or a conflict-aperture interaction ($F(2,24) = 1.01$, $p = \text{NS}$).

**Distortion**

Figure 5.5E shows distortion ratings against decreasing aperture size (increasing DOF and decreasing blur). Results of a 2 x 3 repeated measures ANOVA showed no significant effects of either conflict level ($F(1,12) = 0.13$, $p = \text{NS}$); aperture ($F(1.7,20.9) = 0.03$, $p = \text{NS}$); or of a conflict-aperture interaction ($F(1.7,20.5) = 0.61$, $p = \text{NS}$).

**Clarity**

Figure 5.5F shows clarity judgements (increasing blur) on the y axis against decreasing aperture size (increasing DOF and decreasing blur) on the x axis. A
5.3. RESULTS

Increasing DOF, decreasing image blur with increasing aperture for each of the ratings. Error bars show ± 1 standard error.

2 x 3 repeated measures ANOVA showed no significant effects of either conflict level (F(1,12) = 1.89, p = NS); aperture (F(1.9,23.11) = 0.58, p = NS); or of a conflict-aperture interaction (F(1.7,20.7) = 2.20, p = NS).

Accuracy

Figure 5.6 shows effects of accuracy of all participants, except for those who failed the TNO test. Overall mean accuracy was high: mean absolute difference was 0.89 (SD = 0.90) for Condition 1 and 0.85 (SD = 0.90) for Condition 2. A 2 (conflict) x 3 (DOF) x 3 (depth) repeated measures ANOVA was conducted on mean dot counting accuracy. A significant conflict x aperture interaction was found (F(2.0, 29.4) = 3.86, p < 0.05). There were no other significant effects. Thus there was a different pattern of results depending on conflict level: best accuracy was for appropriate DOF (f11 aperture) in the DOF-vergence only conflict condition (Condition 1). However, accuracy was worst for appropriate DOF (f11 aperture) in the DOF-accommodation-vergence (Condition 2) when all participants with stereovision were included.
5.4 Discussion

There were no significant effects of DOF blur on subjective ratings of photographs. Condition 1 investigated the effects of DOF-vergence conflict only. It was predicted that conflict arising from inappropriate DOF cues (either too much or too little, given human optics) would have resulted in increased discomfort judgements. No effects of increasing DOF on discomfort judgements were found. There was also no main effect on accuracy at the dot-counting task.

Condition 2 investigated the effects of DOF blur on an existing accommodation-vergence conflict such as that typically present in stereoscopic displays. There
5.4. DISCUSSION

was no effect of increasing DOF on discomfort judgements in presentations containing an accommodation-vergence conflict. The accommodation-vergence conflict has already been documented as causing discomfort (Shibata et al., 2011). One of the biggest differences between this current study and that of Shibata et al. (2011) is that the size of the conflict was much smaller in the present study. In the present experiment, the accommodation-vergence conflict was either 0.81°, 1.69° or 2.35° of visual angle, corresponding to 7cm, 17cm, 27cm difference between vergence and accommodation (screen distance). This is 0.22D, 0.46D and 0.64D conflict respectively. In the Shibata et al. (2011) paper, the conflict corresponded to 1.2D, or 0.88m, much larger than the maximum 0.27m conflict in the present experiment.

Typical conflicts in stereoscopic displays work on the principle that the disparity should not correspond to much more than ±55mm in front of or behind the screen, for displays around normal viewing distance (approximately 50 to 100cm viewing distance (Jones, Lee, Holliman, and Ezra, 2001)) (Jones et al., 2001; Holliman, 2004). In the maximum possible conflict level in this range (60mm conflict at 50cm viewing distance) the conflict will be 0.22D. This ‘rule of thumb’ is an extrapolation based on Percival’s zone of comfort. Percival’s zone of comfort is effectively a look-up table of the levels of accommodation-vergence conflict that cause discomfort (and later diplopia with even greater conflicts). As can be seen from these guidelines, the 0.22D conflict should be approximating this conflict level, and the greater conflicts should be outside of the zone of comfort. As the Shibata et al. (2011) study was particularly interested in investigating the discomfort caused by conflicting accommodation-vergence cues, they deliberately set up conflicts much larger than this ‘rule of thumb’ based on Percival’s zone of comfort. The largest conflicts were up to 1.2 dioptres, which corresponded to distances of 10m specified by one cue compared to 0.77m by the other at the furthest distances, and at the closest distances 0.40m and 0.25m cue conflict. Therefore it is perhaps unsurprising that they found discomfort from these highly conflicting stimuli. However, it should be noted that these levels of conflict are not normally those used in stereoscopic presentation (Jones et al., 2001; Holliman, 2004).

In the current study, the DOF blur added corresponded to an even larger conflict than this. For example, in one of the conditions the level of blur was such that the flanking stimuli would have to be at a depth of 8m away, while disparity defined a distance of 53cm (screen distance) or up to 26cm from the observer in Condition 2. Eight metres defined by DOF compared to 53 to 26cm is a larger conflict than the accommodation-vergence conflicts used by
Shibata et al. (2011), however, even this large conflict did not cause discomfort judgements.

It has been shown that DOF blur creates an impression of depth, which could be used to supplement perceived depth from disparity. Thus smaller disparities are needed to create the same impression of depth in observers, by using DOF blur (Wang, Barkowsky, Ricordel, and LeCallet, 2011). It is therefore useful to demonstrate that DOF blur does not cause discomfort, even when large conflicts are present. Therefore perhaps one method of limiting the existing accommodation-vergence conflict usually present from stereoscopic displays is to reduce the disparity, but add DOF blur to create the same percept of depth (Wang et al., 2011), which can be done without fear of causing discomfort from the introduced blur.

The lack of effect of DOF blur manipulations on discomfort is consistent with the idea that accommodation responses are the important factor in determining discomfort judgements. Stimuli not affecting the accommodation response show little effect on discomfort, which seems to indicate that accommodation, rather than blur per se is the critical factor to cause discomfort in the spatial frequency manipulations of the previous experiments (see Chapters 2 and 4). Accuracy showed a differential effect due to the two conflict conditions. In the DOF conflict only condition, most accuracy was with the appropriate (f11) level of DOF blur. However, with the accommodation-vergence condition, accuracy was worst with appropriate levels of DOF blur. Therefore DOF blur might be adding an additional cue to the accommodation-vergence conflict to reduce the conflict. This is possibly due to a reduction in conflict level from blurring the small disparities on the flanking objects, which could reduce the strength of the vergence cue. This effect was not strong enough to be a main effect in its own right, and was only significant when all observers were included in the analysis. This indicates that effects of DOF blur are very subtle in the current sample.

If extreme manipulations are needed to see effects on discomfort judgements, it is possible that these effects and judgements are not strong enough under everyday circumstances to influence the general population. This leads to questions of the importance of the manipulations in causing discomfort. If manipulations that cause discomfort have important consequences for the general population, then if should be the case that they impact performance. Therefore in the next two chapters, the influence of manipulations that report-
5.4. DISCUSSION

edly increase discomfort on performance on visual tasks will be assessed. In Chapter 6, the impact of uncomfortable stimuli on visual search performance will be assessed.
CHAPTER 6

EFFECT OF UNCOMFORTABLE IMAGES ON VISUAL SEARCH

6.1 Introduction

This work was in collaboration with Dr. A. D. F. Clarke of the University of Edinburgh

The focus of the previous chapters was to investigate the properties of uncomfortable images, and to attempt to explain these in terms of the efficiency of neural coding and the effectiveness of the stimuli in driving accommodation responses. One question that remains unanswered here is the extent to which these stimuli have an impact on people’s lives. This can be investigated to some extent in the laboratory in terms of their effect on performance in visual tasks.

Impact of Visual Discomfort

There is existing evidence that visual discomfort could affect visual performance in the general population. Much of the previous evidence concerns individuals who score highly on measures of visual discomfort susceptibility, such as the Visual Discomfort Scale, or VDS (Conlon et al., 1999). Conlon et al. (2001) found that high scorers on this scale have poorer contrast discrimination thresholds than low scorers. Conlon and Humphreys (2001) found that high visual discomfort sufferers showed longer reaction times than controls on a pop-out visual search task, where the target was of a different orientation from the distractors. Additionally, high scorers on the VDS showed more of
a detriment to performance with increasing set size, indicating that the detriment was not a simple overall slowing for this group, but an interaction with the stimulus.

As well as affecting low level visual tasks, there is some evidence that susceptibility to visual discomfort might have an impact on everyday tasks: individuals with high scores on the VDS are associated with lower academic achievement compared to those with low scores (Chase et al., 2009). Additionally, those with high scores on another scale of visual discomfort, the Pattern Glare Test, have also shown a reduced rate of reading compared to low scorers (Allen, Gilchrist, and Hollis, 2008). Although individual differences are not the subject of this study, this research demonstrates that those perceiving more discomfort show poorer visual task performance, in samples drawn from the general population.

As seen in previous research, and in previous chapters, the amount of perceived discomfort experienced from stimuli can be influenced by their spatial attributes. It might be expected that manipulations that increase discomfort, such as changing the spatial frequency content of either noise patterns or striped gratings, would have a detrimental effect on performance in visual tasks.

The global stimulus configuration can be an important factor in discomfort: chequerboard stimuli have been shown to be less problematic than gratings for epilepsy sufferers (Wilkins et al., 1975). Also, a shorter aspect ratio of stripes, making them more like chequerboard patterns, has been shown to produce less perceived discomfort in non-clinical populations (Wilkins et al., 1984). Manipulations of global configuration have also been found to affect performance on visual search tasks. Conlon et al. (1998) investigated reaction times in a task where subjects were instructed to count the number of letter ‘E’s, presented so that their global configuration was plaid-, chequerboard-, or square-wave-like. Observers showed longer reaction times when the stimuli were presented in a square-wave-like pattern compared to the chequerboard- and plaid-like configurations. Additionally, high scorers on the VDS had slower reaction times on a visual search task than low or moderate scorers on this task. These results suggest that the longer reaction times found for square-wave-like configurations result from the visual discomfort associated with these stimuli, and this is more so for those reporting susceptibility to visual discomfort.

Similarly, spatial frequency is another stimulus property affecting visual per-
formance that is also associated with discomfort, both for noise and striped patterns (see Fernandez and Wilkins (2008); Juricevic et al. (2010); Wilkins et al. (1984) as well as Chapters 1 to 4). Spatial frequency manipulations have been shown to affect discrimination of letters hidden against a striped background (Chronicle and Wilkins, 1996). Additionally, Coulon and Hine (2000) used an oblique line target amongst vertical distractors to investigate visual search performance in high and low discomfort groups. Both discomfort groups showed increased reaction time when a striped background was present, the 2 cycles/degree background resulting in a greater detriment to performance than the 15 cycles/degree background. As the 2 cycles/degree stimulus is in the range more associated with visual discomfort than the 15 cycles/degree stimulus (Wilkins et al., 1984), this result could be interpreted as an effect of visual discomfort on reaction times in visual search tasks.

Measures of performance

Visual search was chosen as the task in the current experiment. When measuring performance in a search task care must be taken to investigate all aspects of performance, as an effect on reaction time alone has not been universally found in previous studies investigating the relationship between search performance and visual discomfort. For example, Allen et al. (2008) measured search performance of high scorers on another measure of visual discomfort, the Pattern Related Visual Stress (PRVS) score. They found no difference in reaction time, but there was an increased error rate in the high scoring group. This was using a visual search task that involved counting the number of a particular digit in a window surrounded by an outer background of letters. This suggests that the presence of a periodic background may detrimentally affect search performance, and that it is important to take account of both reaction time and error rate when assessing performance.

Aim

The aim of the current study is to investigate visual search performance against search areas identified by previous research as being comfortable and uncomfortable. It is hypothesised that uncomfortable patterns will detrimentally affect visual search performance compared to more comfortable stimuli. Overall, poorer performance is expected for uncomfortable stimuli compared to comfortable stimuli. More specifically, if stimuli previously identified as uncomfortable are aversive, it might be expected that observers will adopt a search strategy that minimises the time for which they are viewed. Thus the
search might be abandoned more quickly against uncomfortable, compared to comfortable backgrounds. The effects of visual discomfort are particularly relevant to visual tasks that require prolonged viewing, since discomfort tends to build up over time (e.g. Chase et al., 2009). Thus it is hypothesised that increased discomfort would decrease motivation to continue to search for an extended period of time.

6.2 Experiment 1: Filtered Noise Backgrounds

In Experiment 1, $1/f^\beta$ filtered noise patterns were used as the search areas, where $\beta = 1$ has been shown to be judged as more comfortable, and $\beta = 2$ as being less comfortable (Juricevic et al., 2010). Experiment 1 consists of two parts. The first part of the study determines 75% correct contrast detection thresholds in order to provide a performance-based metric to match task difficulty against the two noise search areas, when viewing time was fixed. A short (2s) and a long (10s) presentation time were used, in order to obtain two levels of difficulty for the subsequent search task. The second part of this experiment involved a visual search task against filtered noise search areas (see Figure 6.1 for an example of the stimuli.) There were three target conditions: easy, hard and absent target, and two search area conditions: $\beta = 1$ or 2. The easy target present (TP) condition used the target contrast found for the 2s threshold experiment in the preceding part of the experiment. Similarly, the hard TP condition used the results of the 10s threshold experiment. The critical comparison involves the target absent trials: it was predicted that observers will quit searching faster against uncomfortable search areas compared to comfortable search areas.

Method

Apparatus:
Stimuli were presented on a 20 inch Iiyama HM204DT A Diamontron V3-CRT Vision Master monitor, which was calibrated using a Minolta LS-110 photometer. The refresh rate was 100Hz, and the resolution was 1280 x 1024 pixels. Observers were at a distance of 50cm, meaning that the 1024 x 1024 pixel search background image subtended approximately 31° of visual angle. Luminance range was between 1.4 and 34.7 cd/m². An Eyelink 1000 (SR Research Systems Ltd) eye tracker was used during the visual search task part of Experiment 1. Software from the Eyelink Toolbox Version 1.4.4 (Cornelissen,
6.2. EXPERIMENT 1: FILTERED NOISE BACKGROUNDS

Figure 6.1: Filtered noise search backgrounds

Left: Gabor target against $1/f^2$ background. Right: Gabor target against $1/f$ background. Target contrast is much higher in this demonstration than in the experiment, where it was determined individually for each observer.

Peters, and Palmer, 2002) was used to operate the hardware.

Observers:
Eight observers (mean age = 24.5, SD = 1.20), who were naïve to the purposes of the experiment, participated in the first part of the study. All observers had either normal vision, or were corrected to normal with the use of contact lenses. All experiments were approved by the University of St Andrews Teaching and Research Ethics Committee.

Stimuli:
Stimuli were created using MATLAB 2006b (The Mathworks Ltd). $1/f^\beta$ filtered noise images were created with a $\beta = 1$ or 2 search background. These backgrounds have previously been shown to be comfortable and uncomfortable respectively (Juricevic et al., 2010). They were created by filtering Gaussian white noise in the Fourier frequency domain. Examples of the stimulus backgrounds are presented in Figure 1. The target consisted of a vertical Gabor patch (3 cycles/degree, $\sigma = 5.2$). The target was presented at a random location on the screen, but avoiding a central strip of $2^\circ$, and a $2^\circ$ strip from the extreme edge of the top and bottom of the screen, and a $1^\circ$ strip from the left and the right extremes. Initial fixations were encouraged to be in the centre of the screen by the addition of a white fixation cross in between each trial.

Procedure:
6.2. EXPERIMENT 1: FILTERED NOISE BACKGROUNDS

Contrast Threshold Determination
There is evidence to suggest that there might be differences in eye movements between images with $1/f^2$ and $1/f$ statistics. Eye movements are determined in part by the spatial properties typical of natural scenes (Zhang, Tong, Marks, Shan, and Cottrell, 2008; Najemnik and Geisler, 2009), and observers tend to avoid fixating in luminance patches of low spatial frequency (Tatler et al., 2005). Specifically, differences in eye movements depending on the slope exponent of filtered noise patterns ($\beta$ values) have been shown (Clarke, Green, Chantler, and Emrith, 2008; Clarke, Chantler, and Green, 2009). In order to control for this effect, the contrast of the Gabor target for the $1/f^2$ and $1/f$ search areas was first adjusted so that the two conditions gave rise to visual searches of equivalent difficulty, (based on individual performance). Additionally, eye movements were recorded to see if this could account for any potential findings.

Detection thresholds for the Gabor target against the $1/f^1$ and $1/f^2$ search backgrounds were determined in order to match stimuli for performance-defined difficulty in the visual search task. Stimuli were presented using the psychophysics toolbox (Brainard, 1997; Pelli, 1997; Kleiner et al., 2007) and the PEST staircase procedure (Taylor and Creelman, 1967). Each individual trial commenced with a fixation cross against a mid-grey background. The stimulus consisted of a target presented against the relevant search background for the condition, either $1/f^1$ or $1/f^2$. Participants were asked to locate which side of the screen the target was, and respond using the left or right arrow keys. The Gabor target could appear at a random location within the noise pattern, avoiding the $2^\circ$ central strip, $2^\circ$ at the top and bottom edges, and a $1^\circ$ strip at the left and right edges. The Gabor target was presented for either a ten second (long, easy) search time, or a two second (short, hard) search time. At the end of this period, the screen was replaced with a blank mid-grey screen. Participants were asked to respond as fast and accurately as possible. This resulted in four blocks, long (easy) and short (hard) presentation times, and $1/f^1$ and $1/f^2$ search backgrounds. The background type ($1/f^1$ or $1/f^2$) and presentation time (ten or two seconds) remained constant for the whole block. Noise was regenerated for each trial, thus no two noise patterns were identical, although their statistics were the same throughout the block. Half the observers started with the $1/f^1$ background, the other half with the $1/f^2$ background. The 10s condition was always completed first. In this part of the experiment the target was always present and the observer indicated which side of the screen the target appeared on using the arrow keys.
The contrast of the Gabor target was reduced until threshold levels for each observer for each condition were found using a staircase procedure. Two interleaved staircases tracking the contrast levels needed for 80% and 60% accuracy were conducted. The reversal rule is determined by a probability function that compares the likelihood of the current stimulus level being greater than the tracked probability (in this case 60% or 80% correct responses). The staircase step size was reduced by half every reversal, and doubled on every third consecutive same response. Full details of the PEST procedure can be found in Taylor and Creelman (1967). A minimum of 150 trials were run in the slow presentation condition, and 300 in the fast presentation condition. The staircases were deemed to have converged within this number of trials, based on pilot data. The entire procedure was repeated for the comfortable and the uncomfortable backgrounds, and for the short (2s) and long (10s) presentation times.

Thresholds were obtained by fitting a logistic function using ’psignifit’ (Fründ, Haenel, and Wichmann, 2011). A bootstrapping procedure consisting of 999 simulations was implemented to estimate confidence intervals. The contrast level at which each observer performed at 75% correct detection was recorded as the threshold contrast level. For each observer a total of four thresholds were obtained: a short duration threshold and a long duration threshold, for the comfortable and uncomfortable stimuli. Thresholds from the short and long durations were then used to give ‘easy target’ and ‘hard target’ conditions in the visual search experiment below. Average thresholds across observers can be seen in Figure 6.2. Note the observers were presented with stimuli according to their individual thresholds, not the average.

Visual Search Task
The results of the first part of the experiment were used to determine four stimulus contrast levels for each individual, based on performance, for each of the two search areas (1/f1 or 1/f2), and for each of the task difficulty levels (easy and hard). The observer’s task in this second part of the experiment was to search for a small Gabor target hidden against either 1/f1 or 1/f2 filtered noise backgrounds. The target contrast was set individually for each observer at the 75% correct detection threshold from the first part of the experiment.

Stimuli were again presented at a random location on the screen, avoiding the central strip and the edges. This time, observers indicated whether the
Average thresholds for ‘easy’ (2 second search time) and ‘hard’ (10 second search time) trials, for the two search backgrounds. Errorbars show ± 1 standard error.
target was present or absent using the arrow keys on the keyboard. There were a total of 800 trials: 400 per condition, separated out into blocks of 100 trials each. Trials were generated in random order. Approximately 30% of all trials were easy trials, with high contrast targets. The easy (high contrast target) trials were included for motivation and attention purposes: in low prevalence search tasks there are many more misses than in high prevalence search tasks (Judd, Durand, and Torralba, 2011). Another approximately 30% were hard (low contrast target) trials, with lower contrast targets. The final 40% of trials were target absent trials. A fixation cross was presented on a mid-grey screen before the onset of each trial. Unlike the first part of the experiment, the second part was self-timed to an extent: observers initiated the trial with a button press, after which the stimulus was presented for a maximum of 10 seconds. Observers were instructed to respond as quickly and as accurately as possible, within this time limit.

Results

Analysis
The percentage of correct trials was used as a measure of accuracy. Reaction times were log-transformed to normalise the distribution prior to statistical analysis. Analysis of eye movements was restricted to fixations longer than 0.2s (Salthouse and Ellis, 1980). One observer had to be excluded from the eye movement analysis as eye movements were not recorded properly. However, behavioural results from this individual were still used. Log transforms of the fixation duration and saccade length were used in the statistical analysis. Statistical analysis was conducted using IBM SPSS Statistics Version 19.

Visual Search Performance
The percentage of incorrect responses for each trial type (easy, hard, absent) is plotted on the right hand side of Figure 6.3. Overall mean accuracy was 63.4% (standard deviation 14.7). Mean reaction times are plotted on the left hand side of Figure 6.3.

Bartlett’s Test showed there to be unequal variances ($\chi^2(2) = 30.12$, $p < 0.01$). As expected, results of a 2 (background) x 2 (target difficulty) repeated measures MANOVA showed there to be an effect of target difficulty (easy or hard) only ($F(2,6) = 42.77$, $p < 0.01$). From Figure 6.3 it can be seen that RTs are shorter on easy trials, then hard trials, then target absent trials. There are also fewer incorrect responses for easy trials than for hard and target absent trials. There is no difference between the two background types.
6.2. EXPERIMENT 1: FILTERED NOISE BACKGROUND

To address the specific hypothesis, results of a repeated measures MANOVA showed there to be no effect of background on either reaction time or accuracy for target absent trials ($F(2,6) = 2.06$, $p = \text{NS}$).

![Figure 6.3: Results of Experiment 1](image)

Reaction time (left) and percentage of incorrect responses (right) as a function of the background type. Results are plotted separately for easy, hard, and target-absent trials. Error-bars in this and all other figures show 1 standard error.

**Eye Movements**

The results of two-tailed repeated measures t-tests showed there to be no statistically significant differences between either mean fixation duration ($1/f^2$ mean = 0.49s, $1/f$ mean = 0.53s) or mean saccade length ($1/f^2$ mean = 259.03, $1/f$ mean = 235.43 pixels) between the two search areas. (Fixation duration: $t(6) = -2.69$, NS; Saccade length, $t(6) = 0.48$, NS).

**Discussion**

Experiment 1 was conducted to ascertain whether there were differences in performance in a visual search task against $1/f^2$ and $1/f^1$ backgrounds, which
have been shown in previous research to influence judgements of discomfort (Juricevic et al., 2010). The results of Experiment 1 show no evidence that observers shifted their criteria in order to avoid prolonged exposure to stimuli previously judged to be uncomfortable. As there was no effect of background on the target present trials, this shows that the difficulty matching part of the experiment was successful. As expected, there was an effect of target type: easier trials were found more quickly and more accurately than hard trials. However, for the specific prediction, there was no effect of background on target absent trials, indicating search motivation was not influenced by task difficulty. The hard trials showed near chance accuracy, which could potentially create a floor effect, and account for the lack of difference between the backgrounds for the hard trials. However, as the easy trials did not show this potential floor effect this explanation cannot account for the lack of effects of background on performance. Secondly, no differences were found between the eye movement behaviour against the two search areas.

One possible reason that the type of background had no effect on performance is that the stimuli were simply not sufficiently uncomfortable to elicit detriments to performance. Therefore, the next experiment investigated the effects of striped patterns on performance using the same visual search task.

6.3 Experiment 2: Striped Backgrounds

Previous research has shown striped patterns to be uncomfortable (Wilkins et al., 1984; Wilkins, 1995), and to be capable of inducing epileptiform EEG patterns of activation in epilepsy sufferers (Wilkins et al., 1979), and even epileptic seizures in those with pattern-sensitive epilepsy (Radhakrishnan et al., 2005). Therefore the study was extended to investigate the effects of striped gratings on visual search performance, as these are expected to be more uncomfortable than the noise patterns.

Previous work on both clinical and general populations has shown spatial frequency tuning in the discomfort of striped patterns (Wilkins et al., 1979, 1984). If any effects on accuracy found are due to discomfort, it is expected that they will be similarly spatial frequency tuned.

The search strategy of observers is very task dependent (Boot, Becic, and Kramer, 2009). Therefore, in order to be able to compare between experiments the search task remained the same: the Gabor target was hidden against a fil-
tered noise search area. However, this time the search was limited to a central filtered noise search area with the rest of the display consisting of an outer background. This method also avoided the performance-matching concerns that were evident in Experiment 1.

Method

Apparatus: The apparatus used was the same as in Experiment 1.

Stimuli: The target was located within a central circular patch consisting of 1/f filtered noise. The central circular region of the patch subtended a visual angle of 6.05°, and had a Gaussian smoothed edge ($\sigma = 1.97°$). In condition one, the target was the same 3 cycles/degree Gabor as before, in Condition two, the target was a 0.75 cycles/degree Gabor. The contrast of the target was fixed to be either 20% (hard trials) or 30% (easy trials). Example stimuli can be seen in Figure 6.4. The background outside the central patch was manipulated: either the background was uniform mid-grey, or a vertical sine-wave grating of 0.75, 1.5, or 3 cycles/degree, at 50% contrast. Luminance range was between 1.35 and 34.66 cd/m².

Observers: 24 observers participated in the study, 12 in condition one, 12 in condition two (mean = 22.5, SD = 1.86). All were naïve to the purposes of the experiment.

Procedure: There were four blocks (one for each spatial frequency outer background) of 100 trials each per condition (mid-grey or striped outer background). Five observers in condition one completed the experiment with 200 trials in each block. The rest all completed 100 trials per block. Reducing the number of trials had no effect on the mean reaction times or accuracies unless fewer than 50 trials were incorporated. Stimuli were displayed for a maximum of 10 seconds. Observers were asked to search for the Gabor target and respond as quickly and accurately as possible using the left (target present) or right (target absent) arrow keys, within the time limit. Participants were informed that the target might be absent, and that the target (if present) would always be located in the central noise pattern. Each trial was initiated by the observers themselves using the down arrow.
6.3. EXPERIMENT 2: STRIPED BACKGROUNDS

Figure 6.4: Striped search backgrounds

Example stimuli for Experiment 2. The target Gabor was always in the central 1/f noise patch. The outer background (mid-grey or striped) was manipulated between blocks.

Results

Figure 6.5 shows the results for the 3 cycles/degree target. The left hand side of Figure 6.5 shows mean reaction time plotted against the spatial frequency of the background, for easy, hard and target absent trials. Reaction time tends to increase with trial difficulty (easy trials have the shortest reaction time), and also with spatial frequency. The right hand side of Figure 6.5 shows the mean inaccuracy, against the spatial frequency of the background, for easy, hard and target absent trials. Again, worse performance (more inaccuracy) is shown with increasing spatial frequency, except for the target absent trials. There are, however, very few inaccurate responses for target absent trials overall.

The results of a 3 (trial type) x 4 (background) repeated measures MANOVA showed there to be a significant effect of trial type (easy, hard, absent) on reaction time and accuracy ($F(4,44) = 30.99$, $p < 0.01$), and also a main effect of background spatial frequency on reaction time and accuracy ($F(4,66) = 6.76$, $p < 0.01$). There was also an interaction effect of trial type and background on reaction time and accuracy ($F(12,132) = 7.33$, $p < 0.01$). Univariate follow-up tests showed that there was a significant main effect of trial type on both accuracy and reaction time ($F(2,22) = 53.90$, $p < 0.01$; $F(2,22) = 35.70$, $p < 0.01$). There was a significant main effect of background on both accuracy and reaction time ($F(2,24.4) = 21.0$, $p < 0.01$; $F(3,33) = 4.75$, $p < 0.01$). There was a significant effect of the trial type x background on accuracy ($F(3.2,35.7)$
Figure 6.5: Results with small target

Mean reaction time (left) and percentage inaccuracy (right) against spatial frequency of background with a 3 cycles/degree target.

\[ F(6,66) = 7.78, p < 0.01 \]

Results of post-hoc repeated measures t-tests show that for the 3 cycles/degree target observers are significantly less accurate when the background was also 3 cycles/degree than when it was 0.75 cycles/degree, for the target absent, easy and hard trials respectively (\( t(11) = 3.03, p < 0.01; t(11) = 4.49, p < 0.01; t(11) = 3.54, p < 0.01 \)). There was also a significant difference in the reaction times for easy and hard trials (\( t(11) = 2.55, p < 0.05; t(11) = 2.64, p < 0.05 \)), but no significant difference for the target absent reaction times (\( t(11) = 1.31, \text{NS} \)).

Figure 6.6 shows the results for the 0.75 cycles/degree target. The left hand side of Figure 6.6 shows reaction time, the right hand side shows inaccuracy against spatial frequency of the background for the three levels of task difficulty (easy, hard, absent). There is an apparent effect of trial type on both reaction time and accuracy. There are only a few incorrect target absent responses. There is an effect of spatial frequency, although in comparison with the results...
6.3. EXPERIMENT 2: STRIPED BACKGROUNDS

Figure 6.6: Results with large target

Mean reaction time (left) and percentage inaccuracy (right) against spatial frequency of background with a 0.75 cycles/degree target.

for the 3 cycles/degree target, the spatial frequency tuning has shifted - worst performance is now found for the 0.75 cycles/degree background, both in terms of reaction times and accuracy.

The results of a 3 (trial type) x 4 (background) repeated measures MANOVA showed there to be a significant effect of trial type (easy, hard, absent) on reaction time and accuracy ($F(4,44) = 20.76, p < 0.01$), and also a main effect of background spatial frequency on reaction time and accuracy ($F(6,66) = 13.72, p < 0.01$). There was also an interaction effect of trial type and background on reaction time and accuracy ($F(12,132) = 15.0, p < 0.01$). Univariate follow-up tests showed that there was a significant main effect of trial type on both accuracy ($F(2,22) = 61.34, p < 0.01$; $F(2,22) = 14.37 p < 0.01$). There was a significant main effect of background on both accuracy and reaction time ($F(3,33) = 37.66, p < 0.01$; $F(3,33) = 14.92, p < 0.01$). There was a significant effect of the trial type x background on accuracy ($F(6,66) = 17.92, p < 0.01$) and on reaction time ($F(6,66) = 12.67, p < 0.01$).
Results of post-hoc repeated measures t-tests show that for the 0.75 cycles/degree target observers are both significantly less accurate and slower when the background was also 0.75 cycles/degree than when it was 3 cycles/degree, for the target absent, easy and hard trials respectively ($t(11) = 9.79$, $p < 0.01$; $t(11) = 6.39$, $p < 0.01$; $t(11) = 10.30$, $p < 0.01$). There was also a significant difference for the reaction times ($t(11) = 3.60$, $p < 0.05$; $t(11) = 3.70$, $p < 0.01$; $t(11) = 5.65$, $p < 0.01$). This can be seen in Figure 6.6: tuning for the worst performance in both longest reaction time and most inaccuracy has shifted to 0.75 cycles/degree background.

Discussion

The results of Experiment 2 show that there were significant main effects of trial type and outer background spatial frequency on accuracy and reaction time at finding a Gabor target hidden against a central $1/f$ noise patch. The effect of trial type is expected: performance is increased on high contrast, easy trials. The main effect of background is more interesting, and can be seen in the spatial frequency tuning effects in Figures 6.5 and 6.6. Poorest performance, both in terms of accuracy and reaction time, is seen when the spatial frequency of the target is similar to that of the background.

The presence of any sinusoidal grating background resulted in worse performance than with a mid-grey background. This could be explained as being due to the discomfort that has been previously reported for striped patterns (Wilkins et al., 1984). The most uncomfortable spatial frequencies have been indicated as being around four cycles/degree for striped patterns (Wilkins et al., 1984). Spatial frequencies around three cycles/degree have previously been suggested as those resulting in maximum discomfort (Fernandez and Wilkins, 2008), and spatial frequency tuning was also found in discomfort judgements of filtered noise patterns (O’Hare and Hibbard, 2011). Therefore if discomfort was the reason for the detriment to performance, it is expected that worst performance will be around three cycles/degree. However, the results show that the spatial frequency tuning depends on the spatial frequency of the target, such that worst performance was found when the spatial frequencies of the target and background were matched. This suggests that a large proportion of the decrement in performance results from the similarity between the spatial frequency of the background and the target, rather than the spatial frequency of the background per se. Therefore, the detriment in performance could be due to other effects such as surround suppression: a reduction in perceived contrast of a target caused by the presence of the peripheral background (Yu,
Klein, and Levi, 2001), possibly resulting from receptive field architecture and the effects of inhibitory interneurons (e.g. Webb, Dhruv, Solomon, Tailby, and Lennie, 2005). Surround suppression affects detection of targets (Petrov, Carandini, and McKee, 2005; Petrov, Popple, and McKee, 2007) and also the perceived contrast of targets (Yu et al., 2001). Surround suppression effects are greatest when target and background are of the same spatial frequency and orientation, and increase with eccentricity (Petrov et al., 2005). Maximum effects of surround suppression are seen when the target and the background at around 10% contrast, and are also important in the range of contrasts used in the current experiment (Petrov et al., 2007).

A final experiment was therefore conducted in which observers searched for a horizontal target, in the presence of a vertical background. If the results are due to discomfort rather than surround suppression, then there should be a residual effect when the target is orthogonal to the stripes.

6.4 Experiment 3: Target Orientation

Method

Apparatus, stimuli and procedure were the same as Experiment 2. Only one target was used, a horizontal 3 cycles/degree Gabor patch.

Observers: 13 observers took part in this study. This included two observers, who had previously participated in Experiment 2 and returned to participate in this part of the study. Two experienced psychophysical observers, (LOH and PBH) with corrected to normal vision, and an additional nine other naïve participants participated in the study (mean age = 23.92, SD = 5.01).

Results

A 3 x 4 way repeated measures MANOVA was conducted to assess the effects of the trial type and background on accuracy and noise. Bartlett’s test showed there to be unequal variances ($\chi^2(2) = 23.32, p < 0.01$). The only significant multivariate effect was that of trial type ($F(4,48) = 15.47, p < 0.01$). Therefore there was no effect of background on accuracy or reaction time when the target was orthogonal to the background. This can be seen in Figure 6.7: there is a clear lack of spatial frequency tuning, indicating no effect of background on the reaction times or the accuracy.
6.5. CONCLUSION

Discussion

The results of Experiment 3 showed that the tuning effect found in Experiment 2 disappeared when the target (a Gabor patch) was at an orthogonal orientation to the outer background stripes. This suggests that the tuning found previously can be explained by surround suppression effects alone, and no effect can be attributed to visual discomfort in the current task.

6.5 Conclusion

Experiment 1 investigated the possibility that uncomfortable filtered noise search backgrounds have a detrimental effect on visual search tasks when efforts were made to match stimuli for task difficulty. The critical comparison was the target absent trials - it was thought that observers would quit searching faster against more uncomfortable backgrounds (thought to be $1/f^2$). Results did not support this hypothesis: there was no evidence of a difference between the two search backgrounds in terms of reaction time, accuracy or eye movements.
In Experiment 2, the search task remained the same, but this time an outer, task irrelevant background was introduced: either mid-grey or striped gratings of varying spatial frequency. It was predicted that if there was a substantial amount of discomfort then there will be a performance detriment that should follow the spatial frequency tuning previously shown in the discomfort literature (Wilkins et al., 1984).

The results of Experiment 2 showed a tuning effect on reaction time and accuracy of search for outer backgrounds of different spatial frequencies. There was a reduction in performance with any of the striped outer backgrounds compared to the mid-grey outer background: both longer reaction times and reduced accuracy. This suggests that any modulation in the background causes a detriment to performance. However, performance was worst when the spatial frequency of the target matched that of the background. As this tuning did not follow the theoretically most uncomfortable background spatial frequency, but instead shifted depending on the target, this suggests the effect is unrelated to discomfort, and possibly resulting from alternative processes such as surround suppression (Yu et al., 2001). This explanation was explored in Experiment 3, in which the orientation of the Gabor target was orthogonal to the striped outer background. Under these conditions no effect of background was observed. This suggests that there are no effects of discomfort on visual search performance in the present experiment.

Effects of discomfort on visual tasks have been previously demonstrated in non-clinical populations, by both Conlon et al. (1998) and Chronicle and Wilkins (1996). One important difference between the Conlon et al. (1998) search task and the present study is that observers were searching for targets in a striped background, whereas in the current situation the background and the target were displaced using the filtered noise search area. The Chronicle and Wilkins (1996) task involved detection of a letter directly against the striped background. Spatial frequency of the target was matched to the background. In the present study, the stimuli were displaced from the striped background. The area of the annulus of the striped grating is important for visual discomfort: (Wilkins et al., 1984) showed that the larger the annulus, and the smaller the centre, the more illusions were reported. Thus the lack of an effect on performance in the current study could be because the target was not directly against the stripes, and this would reduce discomfort effects (Wilkins et al., 1984). Therefore discomfort might have been weaker than previous investigations.
In summary, this study shows that theoretically uncomfortable stimuli did not affect performance on a visual search task in a non-clinical population. Thus, if search strategy was altered by differences in motivation due to discomfort, this effect was small in comparison to interference effects. This suggests that, although non-clinical populations can judge images for discomfort (Juricevic et al., 2010), these discomfort effects are not substantial enough to affect performance. More extreme discomfort responses that are powerful enough to manifest themselves in performance measures might be seen in clinical populations, however, investigating the differences between populations is not the objective of this study and thus remains for future research. This is not to say however that such image properties are not in important factor that affects visual performance, and indeed such effects are reported here, and elsewhere. Rather, it is argued that the degree of similarity between the target of interest and the background is an important determinant of the level of disruption of performance that will occur. In many everyday tasks, such as reading, the target and the background will show many visual similarities. This might then be reflected in the disruption to performance in some studies (e.g. Allen et al., 2008).

Although there was no evidence of disruption to performance on the search task, a different task might show other results. For example, a detriment to reading performance has been reported previously from text patterns (e.g. Wilkins et al., 2007; Chase et al., 2009; Tosha et al., 2009). Therefore, it could be the case that these effects of discomfort on performance need time to accumulate, therefore the brief presentation time of the search task might not have allowed for this. Reading is also a more ecologically relevant choice of task. Therefore, the impact of text manipulations, including text stripiness, on reading performance will be assessed in Chapter 7.
CHAPTER 7

EFFECTS OF DISCOMFORT ON READING

7.1 Introduction

The data for Experiment 1 were collected by Miss C. Gordon

The results of Chapter 6 provided no evidence to suggest that stimuli typically reported as uncomfortable showed any effect on performance in a visual search task beyond suppression effects. Reading performance is associated with visual discomfort (Wilkins and Lewis, 1999). In this chapter the effects of stripiness of text stimuli on reading performance will be investigated.

Reading performance has been associated with visual discomfort, both susceptibility of individual observers to visual discomfort (e.g. Chase et al., 2009; Tosha et al., 2009) and the ability of the stimulus to induce discomfort (e.g. Nahar et al., 2007). Self-reports of the reading experience have been used as a method of classifying individuals into high and low discomfort groups (Conlon et al., 1999). Additionally, manipulations that may reduce discomfort might also improve reading performance. For example, using coloured overlays can improve reading performance for some individuals (Wilkins, 1993). Coloured overlays have been reported to reduce discomfort (Tyrrell, Holland, Dennis, and Wilkins, 1995). It is thought that coloured overlays reduce the intensity of particular wavelengths of light that the user finds more problematic. It has previously been shown that some wavelengths of light can be more problematic than others (Main et al., 1997). In this study, migraineurs were shown
to report more discomfort from blue absorption, reddish light than for white light, whereas control subjects reported most discomfort on viewing a high absorption, blueish light source. Thus by changing the spectral content in this way, discomfort might be reduced, and reading unimpeaded.

**Stripes in Text**

Stripes are a stimulus known to potentially cause visual discomfort (e.g. Wilkins et al., 1984). Text could be thought of as a periodic pattern similar to stripes. Manipulations of global stimulus configuration of striped patterns, such as spatial frequency and duty cycle, have been shown to affect discomfort judgements (Wilkins et al., 1984). Manipulations of global configuration of stimuli, in this case rows of the letter ‘E’, have been shown to affect search performance in high discomfort groups (Conlon et al., 1998). Using a mask to isolate one line of text would change the global configuration. The use of masks to isolate only one line of text is another manipulation shown to be successful in improving reading ability in some individuals (Wilkins, 1993). This implies that manipulations of similar properties in stripes can influence both discomfort judgements and reading performance.

It has been shown that distorting the text to reduce stripe regularity affects reading performance in poor readers (Wilkins et al., 2007). The autocorrelation function is the cross-correlation of a signal with itself, as a function of the spatial separation between the samples, see Equation 7.1.

\[
\rho = \sum \left( \frac{(x - \mu_x)(y - \mu_y)}{\sigma_x \sigma_y} \right) \tag{7.1}
\]

Where \(\rho\) is the correlation coefficient, \(x\) is the first sample (image), \(\mu_x\) is the average of the first sample, \(y\) is the second sample (image with spatial offset), \(\mu_y\) is the average of the second sample, \(\sigma_x\) is the standard deviation of the first sample, and \(\sigma_y\) is the standard deviation of the second sample.

The autocorrelation function could be thought of as a measure of how periodic a stimulus is. If stimuli are highly periodic, there will be strong peaks in the autocorrelation function at the relevant spatial offsets (lags) that correspond to the period of the stimulus. In order to distort the text, Wilkins et al. (2007) manipulated it in such a way that the first peak of the autocorrelation function was reduced. Distorting the text to reduce this first peak resulted in a
subtle improvement in both the speed and accuracy of reading in poor readers.

One account of why this improvement might occur is due to large responses in the brain to very regular stimuli. Regular patterns such as stripes have very different amplitude spectra compared to natural images, and this can be detected by infants (Kleiner, 1987). Kleiner (1987) used face and lattice stimuli to create composite images by exchanging the amplitude and phase spectra. Infants showed increased preferential looking for the stimuli with the amplitude spectrum of a face, despite adults considering these stimuli to look more like lattices. The authors suggested this as evidence that the ‘faceness’ of stimuli is conveyed by the amplitude spectrum, rather than the phase information. By disrupting the regularity text patterns might be coded slightly more efficiently. Neurons in the visual cortex are tuned to their own specific spatial frequencies and orientations (see Chapter 3). Groups of neurons responding to the same spatial attributes make up a particular ‘channel’. These channels are selective for spatial frequency and orientation. When viewing natural images, the range of spatial frequency and orientations available will ensure the response is distributed across all channels (Brady and Field, 1997). Highly regular patterns will demand a strong response from only those neurons with similar spatial frequency and orientation tuning preferences, i.e. from particular channels. This could potentially result in an imbalance in the processing channels, and therefore could cause discomfort (for a more detailed discussion of this topic, see Chapter 3).

**Binocular Mismatching**

Another explanation could be an ocular motor one: a problem with binocular mismatching. Humans have two forward-facing eyes, whose views largely overlap. To see with single vision, the brain must combine these two inputs. This involves matching points in one eye’s view of the world to the corresponding points in the other eye’s view. This is called binocular matching, and is useful for aspects of vision such as stereoscopic depth perception. In order to match points in the scenes, it is necessary to establish which points in the two eyes’ views are the same. This is called the ‘correspondance problem’.

It is unclear how the visual system solves this correspondence problem (see pages 41 to 65 in Howard and Rogers (2002)). Although the visual system routinely manages to correctly match corresponding points in the world, this does not always happen. Under certain situations binocular matching fails, and points are mismatched to ones other than the corresponding one. On
viewing repetitive stimuli, this can lead to an illusion of depth misplacement, known as the 'wallpaper effect' (Howard and Rogers, 2002). This tends to happen in situations where adjacent points are similar to the corresponding point, such as patterns of stripes. For example, the first stripe in one eye could be mismatched with the second stripe in the other eye. See Figure 7.1.

![Diagram of successful binocular matching vs. binocular mismatching](image)

**Figure 7.1: The ‘wallpaper effect’**

The wallpaper effect is a case of binocular mismatching. The two eyes views are not combined properly. In this case, the stripe fixated by the right eye is not the same as that of the left. This happens more often with regular patterns, as there is increased similarity of points to match.

**Binocular Control and Reading**

There are also some studies suggesting that binocular mismatching is a problem that could cause visual discomfort, and potentially reading difficulties, in those that are potentially more susceptible to it (e.g. Grisham, Sheppard, and Tran, 1993). Those more susceptible to binocular mismatching might be those displaying problems with aligning the two eyes in the absence of an object, known as ‘phoria’. Problems with vergence control might lead to suppression of one eye’s input (e.g. Tang and Evans, 2007). Alternatively, the observer might confuse inputs from the two eyes, resulting in illusions such as the wallpaper effect.
When there is no coherent stimulus for both eyes to fixate on, many people’s eyes tend to have a resting state that is to look straight ahead. Phoria, a condition when the eyes are not parallel when there is no coherent stimulus to look at. There is evidence to associate exophoria (eyes looking away from each other) with subjective reports of discomfort, and that it increases after working at a computer screen (Nyman et al., 1985). Additionally, Grisham et al. (1993) found that phoria and fixation disparity scores correlate with subjective reports of visual symptoms such as eyestrain, and the skipping of words in a reading task. In contrast to phoria, where the eyes do not have a target to look at, fixation disparity is the slight mismatch between the exact fixation point of the eyes under binocular viewing conditions. Yekta, Pickwell, and Jenkins (1989) provided evidence that reading text at too close a distance increases discomfort, and there is an increase in phoria after a day of reading. This suggests that fatigue effects might build up over prolonged periods of reading. However, in a more recent study Jaschinski (2002) found that although there are more signs of fatigue when forced to read at too close a distance, the preferred distance might well be dependent on the individual.

There are several ocular motor components to reading that are different in poor readers compared to stronger readers, such as more regressions, more variable saccades, longer duration of saccades, poor vergence, unstable fixation, more blurring, and loss of place (for a review see Kulp and Schmidt (1996)). Any of these could account for the differences between strong and weaker readers, and potentially contribute to the reported discomfort. Altogether, these data provide evidence to suggest that there might well be some ocular motor component to visual discomfort, such as binocular mismatching, that has applications in reading, specifically vergence control. In order to test this, reading will be used to assess the effect of a selection of manipulations designed to reduce the chances of binocular mismatching.

**Manipulations of Text**

Direct interventions designed to reduce binocular mismatching have been found to influence reading performance. Occluding one eye was found to improve binocular fixation and reading ability in dyslexic children (Stein, Richardson, and Fowler, 2000). It is thought that the lack of binocular control leads to confusion between the letters of the text. Removing conflicting information by closing one eye therefore facilitates reading.

Occluding one eye is an obvious way of eliminating the issue of binocular
mismatching. However, there are additional manipulations of the text that could influence the possibility of mismatching. Various text manipulations have previously been shown to affect reading performance, and some involve reducing the chances of binocular mismatching.

A specific manipulation that may reduce the probability of binocular mismatching is the use of serifs: little ‘flicks’ in certain fonts. It is thought that ‘sans serif’ fonts (fonts without these ‘flicks’) would be clearer to read as they would create slight differences between the letters (see Figure 7.2), and disrupt the regularity of the spacing between adjacent letters. These additional differences would lower the chances of binocular mismatching. Yager, Aquilante, and Plass (1998) found that there was a benefit of the fonts without serifs on reading speed in low luminance conditions (less than 0.2cd/m$^2$). However, the fonts used differed on two dimensions - presence or absence of serifs, and line thickness of the letter. This finding was also under low luminance conditions only, which might also have impacted acuity. Therefore, whether the effect on reading was due to the serifs is unclear. As other researchers do not find an effect of serifs (e.g Sheedy, Subbaram, Zimmerman, and Hayes, 2005), the importance of serifs is even more debatable. However, Arditi and Cho (2005) found that at a certain text size, the presence of serifs made a tiny improvement to the reading speed of short sentences presented on a computer in very small fonts. However, the sample size in this study was very small - only four participants, two low vision and two normal vision subjects. Also, this effect was very subtle, and the authors concluded it was possibly due to the increased spacing demanded by the serifs rather than the presence of the serifs themselves.

However, not all studies allow firm conclusions to be drawn. For example, Mansefield, Legge, and Bane (1996) found there to be a benefit of reading speed and accuracy with Courier-Bold font compared to Times New Roman font in low vision subjects. However, there are many differences between the two fonts, such as letter width, differences in serifs, and thickness of line, so it is not clear what the important factors are from this study. Additionally, in the second experiment of their study, (Wilkins et al., 2007) showed there to be no difference on reading speed between Times New Roman (with serifs) and Arial (without serifs). In the fourth experiment in this study, Geneva (sans serif) font was found to be read faster than Times New Roman (with serifs). However, presence or absence of serifs is not the sole difference between these fonts, so there may be an effect of font, all other things being equal.
These studies provide mixed and inconclusive indications that some text manipulations affecting the stripiness of text might affect reading performance in some individuals. Work by Jainta, Jaschinski, and Wilkins (2010) showed more direct evidence that stripiness of text affects binocular fixations. The authors concluded that this could in turn affect the reading speed as found by Wilkins et al. (2007). Jainta et al. (2010) measured the autocorrelation of common words, and the fixations of observers reading them. They found an association between the binocular fixation duration and accuracy, and the autocorrelation function. Observers fixated words with a higher first peak in the autocorrelation function for longer, and there was less fixation disparity for these words. This was the case for binocular viewing only - there was no association between monocular fixation duration and first peak in the autocorrelation function. This indicates that binocular fixations need to be more accurate and take longer to achieve for the stripier words. Thus binocular control could account for the reduction in reading speed found by Wilkins et al. (2007).

Although this provides evidence that binocular fixations could account for the findings of Wilkins et al. (2007), this does not eliminate the possibility that the slower reading speed could be accounted for in terms of large neural responses to regular patterns. Manipulations making the striped pattern more irregular could reduce the magnitude of the neural response, and therefore improve reading performance.

Manipulating the autocorrelation function alone does not discriminate between
7.1. INTRODUCTION

these two possibilities, further manipulations are needed. Manipulations such as closing one eye have the effect of removing the binocular matching problem. However, closing one eye also reduces input to the cortex, thus this manipulation cannot discriminate conclusively between the two possible accounts. One possibility is to investigate whether there is an effect of autocorrelation function peak within the monocular condition. Another condition, where the stripiness is manipulated in a manner that does not facilitate binocular mismatching, would also strengthen conclusions. One way of doing this is to change the horizontal line spacing, which has been shown to affect visual discomfort judgements and performance on visual search (Conlon and Humphreys, 2001). Therefore the aim of this study is to establish the difference between the two, to see if there is an effect of discomfort on reading performance.

Reading as a measure of performance

This study is not an investigation of the cognitive processes underlying reading itself, it is a study of visual discomfort using reading performance as a proxy. Reading is multifaceted, ranging from single words on signposts and labels, to large passages of text. It is important to have a working definition of reading in order to be able to measure performance. Previous studies involving reading have used standardised tests, such as the MNRead test (Legge, Ross, Luebker, and LaMay, 1989). This method involves presentation of short sentences on a computer screen or on cards, which the participant reads aloud and the experimenter records mistakes. Presentation time is reduced until mistakes are made and the fastest error-free time is the rate of reading in words per minute. The Group Reading Test is used in schools to classify poor and strong readers (Tyrrell et al., 1995). Another option is the Rate of Reading Test, (Wilkins, Jeans, Pumfrey, and Laskier, 1996), which involves reading a passage consisting of real words in random order, rendering the text meaningless.

As the current study is not designed to isolate subtle effects to probe the processes underlying reading itself, two measures of reading were used. Comprehension of text is one of the goals of the reading task in this study, and so in Experiment 1 a natural reading task with real passages of text was used. Standardised tests are designed to eliminate possible variables such as context effects of words and differences in contextual knowledge of the subject matter. Thus in Experiment 2, a standard test of reading (the Wilkins Rate of Reading Test (Wilkins et al., 1996)) was chosen as a measure of reading. There was no screening of readers in either study, as reading in the general population was the focus of this study, not in a select subset of individuals with reading
7.2. EXPERIMENT 1

difficulties.

7.2 Experiment 1

The study is to investigate whether visual discomfort or binocular mismatching have effects in the general population when reading natural texts. In this study, reading will be defined as reading short paragraphs of text, and measures of reading ability were taken as the speed, accuracy and comprehension of text. The effects of distorting the text autocorrelation function, adding serifs, and viewing the text monocularly and binocularly will be investigated. It is expected that these manipulations should improve reading in those susceptible to binocular mismatching. A different manipulation that will also be used is changing the line spacing. This will change the stripiness of the text, but in the vertical, rather than horizontal direction. Therefore this manipulation will affect global configuration, and reduce stripiness, but not affect binocular mismatching. Therefore it is hypothesised that there may be more of an effect in this condition for those with high visual discomfort rather than low visual discomfort as measured by the Conlon et al. (1999) visual discomfort scale. However, this manipulation should have no effect on reading if binocular mismatching is the cause of the problem.

Method

Stimuli:
Texts were ten passages chosen from ‘Erewhon’, by Samuel Butler (Butler, 1872). These particular passages were chosen as they were free from names of people and places. The results of a pilot study found them to take around a minute to read, although there was some variation between texts (see Table 7.1). To compensate for this, texts were randomised across participants and conditions, and reading speed was measured (in words per minute) as opposed to time taken to read the whole text.

For the baseline condition, words were printed in 12pt Arial font with single line spacing. For the serif condition, words were printed in 12pt Lucinda Bright. For the non-serif comparison, words were printed in 12pt Lucinda Sans, as these were closely matched to the Lucinda Bright for other text properties. See Figure 7.3.
7.2. EXPERIMENT 1

<table>
<thead>
<tr>
<th>Text</th>
<th>Length words</th>
<th>Observer 1 time(s)</th>
<th>Observer 2 time(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>275</td>
<td>1.57</td>
<td>1.09</td>
</tr>
<tr>
<td>2</td>
<td>305</td>
<td>1.47</td>
<td>1.15</td>
</tr>
<tr>
<td>3</td>
<td>265</td>
<td>1.33</td>
<td>1.12</td>
</tr>
<tr>
<td>4</td>
<td>254</td>
<td>1.24</td>
<td>1.06</td>
</tr>
<tr>
<td>5</td>
<td>266</td>
<td>1.03</td>
<td>0.48</td>
</tr>
<tr>
<td>6</td>
<td>250</td>
<td>1.20</td>
<td>0.58</td>
</tr>
<tr>
<td>7</td>
<td>256</td>
<td>1.27</td>
<td>1.04</td>
</tr>
<tr>
<td>8</td>
<td>240</td>
<td>1.17</td>
<td>0.51</td>
</tr>
<tr>
<td>9</td>
<td>265</td>
<td>1.24</td>
<td>1.00</td>
</tr>
<tr>
<td>10</td>
<td>207</td>
<td>1.25</td>
<td>0.58</td>
</tr>
</tbody>
</table>

Table 7.1: Table of text lengths

Results of a pilot study to ascertain reading length of texts in minutes. Due to the variation found, reading speed is quoted in words-per-minute in the subsequent sections.

**Arial font**

**Lucida bright font**

**Lucida sans font**

Figure 7.3: Fonts

The font types used.

In the morphed condition, text was distorted to manipulate the autocorrelation function (stripiness). The basic 12pt Arial font was distorted thus: the letters in the middle of each word were compressed by 1pt, using the ‘condense’ function in Microsoft Word. In the case of words with an even number of letters, the letter(s) slightly to the left of the middle was/were compressed. For words consisting of two to five letters, only one letter was compressed, for words of six or more letters, two letters around the middle of the word were compressed. Words were stretched using the ‘expand’ function at the right hand side by the same number of letters to compensate for overall word length. Single letter words were not manipulated in this way. An exaggerated example of the morphed text, for demonstration purposes, can be seen in Figure 7.4. An example of the actual stimuli can be seen in Figure 7.5.
7.2. EXPERIMENT 1

An example of the morphed text
An example of the unmorphed text

Figure 7.4: Morphed text

This is an exaggerated example of the condensing of the text for demonstration purposes. As the letters were only condensed or expanded by 1pt in the actual experiment, the spacing between the letters remained.

Figure 7.5: Actual text

This is an example of actual stimuli used in the text condensing manipulation.

The autocorrelation function of each line of the morphed text was calculated for different lags according to Equation 7.1. As previously mentioned, the autocorrelation function is a measure of the strength of the correlation of the image with itself at different separations. With no separation, the autocorrelation will be perfect, a coefficient of 1. With increasing separation (lag, in pixels) the autocorrelation coefficient will drop. Periodic stimuli will be characterised by increases in the autocorrelation function as the stripes line up with each other with progressive lags. Therefore the peaks in the autocorrelation function are a way of quantifying the regularity of a pattern. A regular pattern will have stronger peaks in the autocorrelation function.

The effect on the autocorrelation function of the different text manipulations, shown in Figure 7.6. As can be seen in Figure 7.6 there is a reduction in some of the peaks of the autocorrelation for the morphed text compared to
The autocorrelation function of an average of ten samples of text from the book ‘Erewhon’ (Butler, 1872). Lag in pixels (x-axis) is plotted against correlation coefficient (y-axis). Four text manipulations: normal is the baseline Arial 12pt, morphed is the Arial 12pt adjusted by hand, non-serif is the ‘Lucida Sans’ font, and serif is the ‘Lucida Bright’ font.

Figure 7.6: The autocorrelation function of text

The autocorrelation function of an average of ten samples of text from the book ‘Erewhon’ (Butler, 1872). Lag in pixels (x-axis) is plotted against correlation coefficient (y-axis). Four text manipulations: normal is the baseline Arial 12pt, morphed is the Arial 12pt adjusted by hand, non-serif is the ‘Lucida Sans’ font, and serif is the ‘Lucida Bright’ font.

the normal text. It can also be seen that the ‘Lucida Sans’ and ‘Lucida Bright’ fonts both show lower peaks than the ‘Arial’ font. The difference to the autocorrelation between the presence of serifs in the ‘Lucida Bright’ font compared to the absence in the ‘Lucida Bright’ font, shows that there is little effect on the autocorrelation function due to this manipulation. However, the presence of the serifs, whilst having little effect on the autocorrelation function, might still facilitate binocular matching. This is because the serifs cause there to be slight differences in the shapes at the tips of the letters. These differences should help to minimise the confusion - it will be clearer that these letters are different ones and should not be matched. Thus if there were a difference in performance between the serif and sans serif texts, it might be more explicable in terms of binocular mismatching rather than coding efficiency.
7.2. EXPERIMENT 1

each line. Spacing varied between 1 and 10pt, mean line spacing was 5.06pt, standard deviation was 0.22pts. These manipulations were done using Microsoft Word.

Observers:
Observers were 31 students with corrected-to-normal vision. Participants were reimbursed five pounds per hour for their time. The experiment was approved by the University ethics committee (UTREC).

Procedure: Firstly, standard vision tests were carried out. The test battery included the TNO test of stereoacuity (Nieuwegein, 1972), the Maddox Rod test, the Conlon Visual Discomfort Scale (Conlon et al., 1999) and a custom-made test of reading acuity see Figure 7.7. The TNO test was to ascertain whether the observers had good stereovision. It is thought that difficulties with binocular matching might result in poor stereoacuity, therefore this could be a useful indicator. The Maddox Rod Test is a measure of phoria, another potential indicator of susceptibility to binocular mismatching. Reading acuity was also measured using a custom-made text sheet - the lowest font size legible was recorded at a reading distance of 40cm. The largest font size was 12pt, the smallest font size was 6pt. Texts were again taken from (Butler, 1872) see Figure 7.7.

Participants then read each text aloud, which was recorded onto a tape recorder. There were ten conditions: two viewing conditions (binocular and monocular) and five text manipulations (normal, morphed, with serifs, without serifs, manipulated line spacing). Order of each condition was randomised for each observer. In the monocular viewing conditions, one eye was occluded using an eye-patch. Observers could choose which eye was occluded, but were asked to remain with this choice for all the monocularly presented conditions.

Participants were allowed to view the texts at their preferred reading distance. Participants were not given instructions regarding speed or accuracy at which they should read, but encouraged to progress at their usual rate.

Reading speed was defined as the number of words per minute, and reading accuracy was noted by comparing mistakes on the tape to the printed text. Mistakes were defined as: hesitation (over 1 second), mispronunciation of words, misreading words, missing words out, adding extra words not in the
7.2. EXPERIMENT 1

Let us be grateful to the mirror for revealing to us our appearance only.

Exploring is delightful to look forward to and back upon, but it is not comfortable at the time.

They could not understand how these things could be; my clothes also seemed quite beyond them.

I could not help speculating upon what might lie farther up the river and behind the second range.

The eminent man reassured him with a few cheering words, and then proceeded to make a more careful diagnosis of the case.

Everything was soaking wet and we were half-perished with cold; indeed we were very uncomfortable.

Mention but the word divinity, and our sense of the divine is dulled.

Figure 7.7: Acuity test for reading task

Custom-made acuity test with sentences chosen from the same book as the experiment (Butler, 1872), but not part of the test material.

printed text. Two independent raters judged the tapes for mistakes. Comprehension was tested by answering five questions about the text after reading.

Results

No observers were excluded due to failing the visual acuity test: all participants were able to read the smallest font size (6pt) presented. No observer was excluded due to failing to comprehend the text: the lowest comprehension score was 11, and the mean number of answers correct was 23.4 (SD = 6.7) out of a possible 40, meaning that individuals had understood some aspects of the texts. Inter-rater reliability was high ($\rho(130) = 0.79 \ p < 0.05$). It was not possible to create a high and low discomfort group according to the guidelines of the VDS (Conlon et al., 1999), as too few participants scored highly on this scale - only 4 scored highly enough to be classified as moderate discomfort scorers, and there were no high visual discomfort scorers. Participants were divided into a higher and lower discomfort group after the results of a median split, based on the results of the VDS. Median discomfort of the higher visual discomfort group was 28.1%, and 6.5% of the low visual discomfort group. A score of 35% is needed to be classified as a moderate discomfort sufferer according to the questionnaire. Viewing condition was either monocular or binocular. Text
7.2. EXPERIMENT 1

manipulations were: normal (baseline), morphed text, with serifs, without serifs, line spacing manipulated. Performance in terms of reading speed in words-per-minute and number of mistakes were analysed using a mixed Design ANOVA, with 2 (viewing condition) x 5 (text manipulation) repeated measures factors, and one between subjects factor (discomfort group).

Reading Accuracy

Figure 7.8 shows there to be a clear difference between the two discomfort groups: the high discomfort group (indicated in red) show more mistakes over all the texts than the low visual discomfort group (blue).

Figure 7.8: Number of mistakes when reading manipulated text

Mean number of mistakes for each of the five text manipulations under the two viewing conditions (binocular and monocular). The high visual discomfort group (red) and the low visual discomfort group (blue) according to the Conlon et al. (1999) VDS are indicated. Error bars show ± 1 standard error.

Results of a mixed design ANOVA involving 2 (viewing condition) by 5 (text manipulation) repeated factors and a between subjects factor (discomfort group) showed there to be no effect of viewing condition (F(1,30) = 1.61, p = NS), or
of text manipulation ($F(2.8, 84.4) = 0.20, p = \text{NS}$). There was no significant interaction between text manipulation and group ($F(2.8, 84.4) = 0.22, p = \text{NS}$), or between viewing condition and group ($F(1, 30) = 0.01, p = \text{NS}$), or between viewing condition and text manipulation ($F(3.5, 105.1) = 1.93, p = \text{NS}$). However, there was a significant effect of group ($F(1, 30) = 8.59, p < 0.01$).

**Reading Speed**

Figure 7.9 shows average reading speed against text manipulation for two viewing conditions (monocular and binocular). Again, the red bars indicate the high visual discomfort group, the blue bars indicate the low visual discomfort group, according to the Conlon et al. (1999) VDS. Again, the low visual discomfort group show faster reading speed than the high visual discomfort group, both under binocular ($t(158) = 4.50, p < 0.01$) and monocular ($t(158) = 4.52, p < 0.01$) viewing conditions.

![Figure 7.9: Average reading speed of manipulated text](image)

Mean reading speed (wpm) for each of the five text manipulations under the two viewing conditions (binocular and monocular). The high visual discomfort group (red) and the low visual discomfort group (blue) according to the Conlon et al. (1999) VDS are indicated. Error bars show ± 1 standard error.
7.2. EXPERIMENT 1

The results of a 2 x 5 x 2 mixed design ANOVA with discomfort group as the between subjects factor showed there to be a significant effect of viewing condition only on reading speed (F(1,30) = 6.90, p < 0.05). There was also a significant effect of group (F(1,30) = 4.37, p < 0.05). From Figure 7.9 it can be seen that the binocularly viewed condition was read the fastest.

Correlations

Average reading speed across conditions for each individual was calculated. Reading speed did not correlate with horizontal phoria. There was a significant negative correlation of reading speed with visual discomfort score: the higher the Conlon et al. (1999) VDS score, the slower the reading speed overall (\(\rho(32) = -0.51, p < 0.01\)). There was a significant positive correlation between reading speed and comprehension (\(\rho(32) = 0.41, p < 0.05\)). This indicates that those who read faster answered more of the comprehension questions correctly.

Discussion

Experiment 1 showed there to be an effect of discomfort group on reading performance: the low discomfort group read both faster and with fewer mistakes. Correlations also showed that those who reported more discomfort tended to read more slowly. This is possibly because the Conlon et al. (1999) Visual Discomfort Scale consists of many items related to reading: only three out of the 23 questions are unrelated to reading.

Those who read faster also answered more comprehension questions correctly. This could be an indication that these individuals are simply better at reading than others. This also means that the increased speed of reading is unlikely to be at a cost of comprehension, therefore individuals are reading quickly whilst still extracting relevant information from the text.

There was no correlation between speed of reading and horizontal phoria. This does not provide any support for the idea that such binocular problems account for the difference in performance between high and low discomfort groups in the current sample.

The results of Experiment 1 showed little effect of text manipulation on read-
7.2. EXPERIMENT 1

Figure 7.10: Correlations between reading speed, discomfort and comprehension.

Left: correlation between average reading speed (wpm) across all texts and discomfort score according to the VDS. Right: correlation between average reading speed (wpm) across all texts and number of comprehension questions answered correctly.

Aside from the differences in the calculation of the autocorrelation function, there are also other methodological differences between the Experiment 1 and the work of Wilkins et al. (2007). Therefore, a standardised measure of reading...
performance was used in Experiment 2, to be able to compare to the study by Wilkins et al. (2007) more directly.

7.3 Experiment 2

Experiment Two was conducted as a more direct replication of Wilkins et al. (2007), to ascertain whether the lack of text manipulation effect was due to the reading task set.

Method

Stimuli:
22 monosyllabic words of either four or five letters long were chosen from the 100 most frequent in the English language (based on the Oxford English Corpus). They were printed on a calibrated hp LaserJet 1200 printer, which prints at 1200 dpi. The words were presented in paragraphs of 300 or 310 words, in approximately 18 lines of single spaced text. Spacing between the lines was 1.15cm.

Words were divided into two groups, more stripy and less stripy, based on autocorrelation function of each individual word. The autocorrelation function of words in Times New Roman 12pt font was measured, see Equation 7.1. Stripy words were defined as those with an above-median first peak in the autocorrelation function. The average (mean) autocorrelation coefficient of the stripy words was $\rho(11) = 0.95$, SD = 0.010. Less stripy words were those with a lower first peak in the autocorrelation function: mean coefficient was $\rho(11) = 0.92$, SD = 0.006.

The autocorrelation function was also measured for the Arial 11.5pt font. This showed some variation compared to the initial analysis using Times New Roman 12pt analysis. The average first peak values of the more and less stripy groups differed: the Arial analysis showed the mean correlation coefficient of the first peak to be $\rho = 0.98$, SD = 0.003, and $\rho = 0.96$, SD = 0.002, for the more and less stripy groups respectively.

This suggests that overall the Arial is more stripy than the Times, and there is less variation between the words. This is to be expected as these are different fonts. Critically, there was little change between the words making up the two groups; four out of the 22 words swapped groups under the Arial analysis compared to the Times analysis. It was considered more important in the
current study to keep the word lists the same across conditions, as changing them could have introduced potentially more problematic biases, such as time needed to articulate words etc. Therefore, as the lists were fairly similar for both of the two analyses, it was decided to just retain the one word list, rather than risk introducing other biases. As the variation was more pronounced under the Times analysis, this list was chosen.

There were three manipulations in this study; autocorrelation function peak (more or less stripy text), viewing condition (monocular or binocular), and font (Times New Roman 12pt, Arial 11.5pt).

Observers:
20 psychology students with normal or corrected to normal vision participated in the study. Ages ranged from 18 to 30.

Procedure:
Participants viewed the stimuli at a distance of approximately 40 cm. No aide to reading was permitted, such as pointing, using a ruler etc. Observers read aloud as many words as possible in one minute into a cassette tape recorder. Mistakes were defined as words transposed, omitted, or incorrectly read, after Wilkins et al. (2007). Under binocular conditions, participants viewed stimuli with both eyes, under monocular conditions observers were free to choose which eye they wished to use, but asked to remain with their decision in subsequent monocular conditions. A short pause ensured that the eye recovered to read the next text. Stimuli were presented in pseudo-random order; all stimuli were the first at some point, and stimuli alternated between monocular and binocular viewing conditions, to ensure that they covered eye did not become too dark-adapted. A practise text of three-letter words, as used by Wilkins et al. (2007) was used to familiarise participants with the procedure prior to commencing the actual experiment.

Results
Two observers had to be excluded from the analysis: one for failure to comply with instructions, the other as they reported vision in only one eye. Mistakes were low (mean = 2.1 per text, SD = 1.3), indicating that observers remained fairly accurate throughout the texts. Therefore any differences in reading speed are unlikely to be due to a speed-accuracy trade-off.

Results were analysed using a 2 (viewing condition) by 2 (font type) by 2
(text stripiness) repeated measures ANOVA. Viewing condition was defined as monocular or binocular, font type was either Arial or Times New Roman, and text stripiness was more or less stripy, based on the autocorrelation function analysis (for details see stimuli section). Reading speed was measured in words-per-minute (wpm). Accuracy was number of mistakes made.

The results of the 2 x 2 x 2 repeated measures ANOVA showed there to be no main effects of stripiness, font or viewing condition on the speed of reading (defined as number of words read per minute (wpm)). However, there was a significant interaction between stripiness of text and font; repeated measures ANOVA (F(1,17) = 8.50, p < 0.05). This can be seen in Figure 7.11: reading speed (number of words per minute) is plotted against stripiness of text (higher or lower autocorrelation function peak) for the two different fonts (Times new Roman 12pt, Arial 11.5pt). Post-hoc t-test showed there to be a faster reading speed with the more stripy text in Arial 11.5pt font (t(17) = 2.16, p < 0.05), and a faster reading speed for less stripy text in Times New Roman font (t(17) = 2.67, p < 0.05).

Post-hoc tests revealed reading speed to be higher with Arial font compared to Times font when the text was not manipulated (t(35) = 2.64, p < 0.05). However, the effect in the manipulated text condition was not significant, although there was a trend in the opposite direction (t(35) = 1.69, p = NS). Additionally, the less stripy Arial font text was read significantly faster than the manipulated version (t(35)= 2.29, < 0.05).

Discussion

The results of Experiment 2 partially replicated the results of Wilkins et al. (2007), that autocorrelation function affects reading speed, for fonts with serifs only. For serif fonts (Times New Roman 12pt), stripier texts led to slower reading speeds than less stripy words. However, the reverse was found for the non-serif (Arial 11.5pt) font: faster reading speeds were with the striper text compared to the less stripey text. As this was still the case under monocular viewing conditions, binocular mismatching cannot account for the effect. Additionally, as there was no effect of viewing condition, this also leaves no evidence to suggest that binocular mismatching is the cause of any discomfort in
Figure 7.11: Reading speed (wpm) against text stripiness for two fonts.

Interaction effects of font and stripiness of text. Reading speed (wpm) for more (squares) and less (circles) stripey text in either Times New Roman 12 pt and Arial 11.5pt font is plotted. Faster reading occurs for less stripey serif (Times) font and for more stripey sans-serif font (Arial).
this experiment. Additionally, as there were few mistakes made overall, speed-accuracy tradeoff cannot account for the findings. The pseudo-randomised order means that there the possibility of practise effects is accounting for the findings is minimal.

It could also be argued that this could be a slight bias in the method of categorising the words as either more or less stripy. The individual words were categorised based on their autocorrelation function measured as calculated in 12pt Times New Roman font. This possibly introduced a bias as some (four) of the words shifted group when the autocorrelation function was measured in Arial font. However, fewer than half the stimulus words moved group when analysed in Arial font. If this were a source of bias, it would be expected to add noise, not reverse the effect. A better method of comparing fonts would have been to calculate the autocorrelation of a large sample consisting of same words in both fonts, and compare the average peak in the autocorrelation function. Therefore, it is not possible from this experiment to draw conclusions about the suitability of certain fonts themselves.

It could be argued that the effect with the Times 12pt font can be accounted for in terms of larger cortical responses: increased text regularity might have caused a higher response in a particular spatial frequency channel, and thus increased metabolic expenditure (see Chapter 3 for a more detailed discussion of sparse coding models). This might therefore have caused discomfort. However, it is difficult to account for why there would be faster reading with stripier text in the Arial condition, therefore this account alone is insufficient.

Binocular control problems could provide an alternative explanation to account for some of the findings. Jainta et al. (2010) showed that there was an increase in accuracy for the binocular fixations with the Times 12pt font. Observers were slower to read words with a higher first peak in the autocorrelation function. However, the binocular mismatching explanation cannot account for the results with the Arial 11pt font. As this font had a higher overall peak in the autocorrelation function than the Times 12pt, it might be expected that this would show an exaggeration of the findings with the Times 12pt font. However, the opposite result is found. Also, there was no difference between the monocular and binocular viewing conditions, which would be expected if binocular mismatching were the cause of the reading speed differences. Font effects on reading speed were also found by Wilkins et al. (2007), in the case of the Geneva (sans serif) compared to the Times New Roman (serif) font. The Geneva font was read more slowly than the Times New Roman. The authors
ascribed this not to the serifs, but to the difference in autocorrelation function between the two font types. However, in contrast to this finding, and to the results of the current experiment, these authors found no difference between the Arial and the Times New Roman fonts. Due to the mixed results, it seems likely that the differences in reading speed found for font are for reasons other than the presence or absence of serifs.

There are other, alternative explanations aside from binocular mismatching and large neural responses. There might be other ocular motor control issues to account for this finding, like saccade planning and execution, as there have been differences found in saccades of weak readers (Eden, Stein, Wood, and Wood, 1994). However, the aim of the present study is not to investigate the processes behind reading, but to use it as a proxy for investigating the effects of stripiness on the visual system using a stimulus and task that are applicable to everyday life, namely text. The two explanations proposed could not account for the data, and therefore future work is needed to ascertain the reasons behind this interaction. Additionally, the effects found by Wilkins et al. (2007) are not limited to reading alone - faster search times were found with less stripey text patterns. This indicates that the problem of stripiness of words cannot be accounted for by language ability and mechanisms of reading per se.

Contrast is another attribute that affects reading, and also affects discomfort. However, effects of contrast are not simple: they vary depending on the type of stimulus. For example, low contrast text (Nahar et al., 2007), but high contrast striped patterns (Wilkins et al., 1984), increase discomfort judgements. Therefore, the effects of contrast on striped patterns and natural images will be investigated in the next chapter.
CHAPTER 8

EFFECTS OF CONTRAST AND LUMINANCE ON VISUAL DISCOMFORT

8.1 Introduction

This work was in collaboration with C. Foubister, an undergraduate project student, and P. N. Day, J. Colville and C. Rohan, of NCR Corp., Dundee. The data for Experiment 2 were collected by C. Foubister, and separately analysed and submitted as part of her undergraduate degree.

Contrast is an attribute that plays an important role in discomfort judgements (e.g. Wilkins et al., 1984; Fernandez and Wilkins, 2008; Juricevic et al., 2010). The effect of relative contrast amplitude at various spatial frequencies has already been previously investigated (see Chapters 2 and 4). In this chapter the level of contrast and the level of luminance will both be systematically investigated. Luminance is the overall level of light reflected, measured in cd/m². Contrast is the variation in luminance over an area, which could be measured in a variety of ways, such as RMS, Michelson (difference between the light and dark regions) or band-limited contrast (see (Peli, 1990)). It is possible that these two stimulus attributes have different effects on discomfort judgements. This possibility will be investigated in this chapter.
8.1. INTRODUCTION

The Role of Contrast

Adverse reactions to visual stimuli are affected by contrast, both in clinical and non-clinical populations. Increasing the contrast of square-wave gratings has been shown to increase the probability of paroxysmal EEG activity in epilepsy (Wilkins et al., 1979). It has been shown that increasing the overall luminance of a visual stimulus increases the likelihood of eliciting a migraine attack in those susceptible (Drummond, 2006). Discomfort judgements made by non-clinical populations are also affected by contrast, both in gratings (Wilkins et al., 1984) and in more complex patterns such as filtered noise (Fernandez and Wilkins, 2008). In the case of grating stimuli, increasing contrast increases discomfort judgements (Wilkins et al., 1984). A similar pattern exists in more complicated stimuli: higher contrast filtered noise patterns are also more likely to be judged more uncomfortable than lower contrast noise (Fernandez and Wilkins, 2008).

It has been argued that excessive cortical responses might be the cause of discomfort in some situations (Juricevic et al., 2010), see Chapter 3 for a detailed discussion. It could be argued that any high luminance or high contrast pattern could potentially cause large response amplitude in the visual system, and should therefore be potentially problematic. In order to determine exactly which attribute causes the excessive response, a distinction must first be made between overall mean luminance and contrast. High luminance is a large amount of light emitted by the stimulus. Increased luminance could induce increased responses in terms of firing rate of neurons (Clarke, Zhang, and Gamlin, 2003). However, gain control mechanisms in the visual system are in place to reduce the sensitivity of the visual system when faced with high overall luminance levels: for example, physiological changes in pupil size restrict the amount of light entering the eye. There are also cortical gain control mechanisms in place that reduce the perceived brightness of light when adapted to this level (Craik, 1940). This is a very rapid adaptation process, taking about 1/20th of a second. It is also effective for large discrepancies in light levels: a dark adapted eye shown approximately 10.5 cd/m\(^2\) can be perceived to be the same brightness as an unadapted eye shown 52,500 to 262,500 cd/m\(^2\) light (Craik, 1940).

A large contrast range could also cause large cortical responses and therefore be uncomfortable (e.g. Juricevic et al., 2010). However, as with luminance, there are also cortical gain control mechanisms for contrast (e.g. Bonds, 1993). Contrast normalisation is one possible gain control mechanism, which is a
non-linear transform to reduce the effect of high contrast ranges (see Tolhurst and Heeger (1997)). Large differences in contrast might be more problematic than luminance, as the contrast gain control mechanisms do not operate for such a large discrepancy: response of striate cells saturates to the maximum firing rate at around 30% contrast of a mean luminance level of 27.4 cd/m$^2$ (Albrecht and Hamilton, 1982). There has been evidence to support the suggestion that lack of cortical gain control processes in epilepsy sufferers might be a contributing factor to their typically excessive cortical activity (Porciatti, Bonanni, Fiorentini, and Guerrini, 2000).

**Discomfort from Low Contrast Stimuli**

However, it is not always the case that increasing contrast leads to increased discomfort: for example, increasing contrast of text was found to lower discomfort reported after completing a reading task (Nahar et al., 2007). This cannot be explained in terms of overly large cortical responses. Instead, this discrepancy between findings using striped patterns and those from text stimuli might be explained as being due to differing task demands. In the case of text, observers are actively trying to extract information from the text in order to read it. To do this, observers need sufficient information to drive the relevant ocular motor processes, for example accommodation and vergence eye movements. These ocular motor processes are also affected by stimulus attributes, notably contrast; for example, it has been shown that a higher contrast stimulus provides a more informative signal to the accommodation system than a lower one (Kotaluk and Schor, 1987). As the observer is actively trying to accommodate this stimulus accurately in order to read it, rather than just passively view the image, it could be that higher contrast is beneficial to provide enough information to facilitate ocular motor processes.

Contrast effects from studies involving striped patterns show that discomfort from gratings increases with increasing contrast, while studies involving text show that discomfort decreases with increasing contrast. The dichotomy between contrast effects for stripes and text could also be explained by differences in stimuli and procedure: different conditions, stimuli, contrast ranges and observers were employed by Wilkins et al. (1984) compared to Nahar et al. (2007). Firstly, one difference was that the task by Nahar et al. (2007) was conducted on a computer screen, whilst Wilkins et al. (1984) presented printed copies of their stimuli. Visual display units (VDUs) have been reported to affect discomfort themselves (Lovasik and Kergoat, 2008). Secondly, different contrast ranges were employed: the stimuli in Nahar et al. (2007)
ranged in contrast from 5% to 40%, whereas the Wilkins et al. (1984) stripes ranged from around 1 to 60% contrast. Thirdly, overall luminance was also not comparable between the two studies; the Wilkins et al. (1984) striped stimuli were around 100 cd/m\(^2\), whereas the Nahar et al. (2007) stimuli were considerably brighter, varying between 817 cd/m\(^2\) to 3388 cd/m\(^2\). Overall luminance could be a critical factor in discomfort, rather than contrast *per se*. Due to all these crucial differences, as well as other, more subtle differences, it is not fair to compare the two studies directly. It would be beneficial to see if these results of increasing contrast being uncomfortable for stripes, and lower contrast for being uncomfortable text, can be replicated under directly comparable conditions, so that a more direct comparison could be made.

**Display Technology**

Display screens are a ubiquitous aspect of modern life. One of the major challenges faced in developing displays is the need for them to be easy to use across a wide range of viewing conditions. While standard emissive displays work well under moderate lighting conditions, as found in the typical office environment, significant challenges are faced in creating displays that are legible and comfortable to use under conditions such as direct sunlight. This can create not only problems of glare, reflections and haze across the screen, but can also markedly reduce the effective contrast and visibility of the displayed image. Two important factors contribute to this problem. The first is the fact that ambient light will be reflected from the display screen, thus reducing the actual contrast of the displayed screen image. While this can be reduced by the use of anti-reflective coatings, these coatings do not entirely eliminate the problem of reflected light and so there is still a reduction in contrast. The second is that, because daylight luminance can be many times brighter than the luminance of a typical display, the visual system will adapt to this high level of ambient lighting, therefore becoming less sensitive to the luminance levels presented on the display (Rushton, 1965). These problems have been considered in relation to the design of displays for self-service terminals (Day, Colville, and Rohan, 2010) and electronic books (Lin, Lin, Hwang, Jend, and Liao, 2008).

High-bright displays present one possible solution to the problem of daylight viewing of screens. High-bright displays have a backlight that can emit in the order of 1700-2000 cd/m\(^2\), making them more useful than traditional displays in bright daylight. These displays have been demonstrated to aid legibility of on-screen content under direct sunlight (Day et al., 2010).
Aim

The aim of the current set of experiments is to investigate the effects of contrast range and overall luminance level on discomfort and clarity judgements. This was done using both intrinsically comfortable (natural images) and uncomfortable (gratings) stimuli. The magnitude of contrast range and mean luminance were both manipulated. It is predicted that overall luminance should not affect discomfort judgements, as gain control mechanisms should reduce the sensitivity of the visual system to these stimuli. However, this will not be the case for high contrast patterns, and therefore it is predicted that higher contrast patterns should be more uncomfortable than lower contrast patterns, even when they are potentially less bright in terms of mean luminance levels.

8.2 Experiment 1

Experiment 1 was conducted to investigate the effects of contrast on discomfort and clarity judgements for text and striped patterns.

Method

Stimuli:
The two stimuli used were created using Matlab (The Mathworks, Inc, 2005). Stimuli were printed on A4 white paper on a HP Laserjet 1200 printer (calibrated using a Minolta LS-110 photometer). Stimuli were chosen to be as similar as possible; the two stimuli used were horizontal stripes and text patterns, measuring 20 x 15.5cm. Viewing distance was approximately 1m, determined by a metre rule. Stripes were sine waves with a spatial frequency of approximately 3.5 cycles/degree at this viewing distance. Text was printed using Arial 14pt font. The stimuli consisted of five strips of either text or stripes, varying in Michelson contrast (see Figure 8.1). Michelson contrast was defined as the proportion of the maximum contrast range, and this was manipulated so that both stimuli had Michelson contrasts of 1.0, 0.80, 0.60, 0.40 and 0.20. Ambient lighting against the white paper background was measured to be 38.22 cd/m².

Observers:
8.2. EXPERIMENT 1

Figure 8.1: Stimuli for Experiment 1

Contrast varied from top to bottom for (left) striped gratings, and (right) text stimuli.

22 university students with corrected to normal vision participated in the study. Age ranged between 18 and 30 years. There was a mixture of both male and female participants.

Procedure:
Observers stood approximately 1m from the wall where the stimuli were positioned. Stripes were on the left, text on the right. Participants were asked to rank all of the stripes from the most to least uncomfortable. They were then asked to rank the stripes from most to least clear. The procedure was then repeated for the text patterns. Stimuli were identified by numbers. Answers were reported verbally and recorded by the experimenter. All observers completed the task in the same order.

Results

The left side of Figure 8.2 plots rankings against increasing contrast for discomfort and clarity judgements. The results of a Friedman’s test showed there to be a significant effect of contrast on clarity judgements in striped patterns ($\chi^2(4) = 66.6, N = 22, p < 0.01$). There was no effect of contrast on discomfort judgements in striped patterns. Therefore lower contrast leads to a trend towards lower clarity rankings, but does not result in increased comfort rankings.

There was an effect of contrast on discomfort and clarity judgements in text
8.2. EXPERIMENT 1

Figure 8.2: Comfort and clarity ratings against contrast.

Left: stripe comfort and clarity ratings plotted against contrast. Right: text comfort and clarity ratings against contrast. Errorbars are ± 1 standard error.

stimuli. This can be seen on the right side of Figure 8.2. There was an effect of contrast on clarity judgements \( (\chi^2(4) = 18.3, N = 22, p < 0.01) \). There was also a significant effect of contrast on discomfort judgements \( (\chi^2(4) = 42.4, N = 22, p < 0.01) \). The result of a Spearman’s correlation shows there to be a significant positive correlation between these two rankings \( (\rho = 0.207, N = 110, p < 0.05) \). This is different behaviour compared to the striped stimuli. This shows that decreasing contrast results in both lower clarity and lower comfort rankings.

**Discussion**

Increasing image contrast increases clarity judgements for both striped stimuli and for text stimuli. Increasing contrast decreases discomfort judgements for text stimuli. However, there is no effect of contrast on discomfort judgements for the stripe stimuli.

The differential effects of contrast depending on the type of stimulus might be because of the way in which observers engage with the stimulus. Observers might find themselves attempting to read the text, and in this case clarity
matters. Reading text is a fairly automatic process for practised readers, and observers might not be able to prevent it (Brown, Joneleit, Robinson, and Brown, 2002).

Reading has some characteristic eye movements (see (Kulp and Schmidt, 1996), for a review). Viewing natural images has a very different pattern of eye movements (Martin and Tadmour, 2004). This is one of many potential differences between viewing regular patterns such as stripes and text, and natural images such as scenes. It was thought that by using natural images there might be a more natural distribution of eye movements than text. Therefore it will be ascertained whether these contrast effects found for text and gratings remain for images with more natural statistics, such as pictures. Therefore in the next experiment, natural scenes and striped patterns will act as stimuli. This means that the highly unnatural eye movements characteristic of reading passages of text will no longer be a confounding variable.

As Experiment 1 was based on rankings within each class of stimulus, this does not allow for meaningful comparisons across the two sets of stimuli. It was therefore decided to conduct an experiment investigating a wider range of stimuli which were divided into two classes.

It could be the case that discomfort effects only occur for higher overall luminances than those in the current study. Although this previous work had also been conducted using printed stimuli (Wilkins et al., 1984), ambient lighting might have meant that overall luminance was relatively high. Moreover, as there are separate gain control mechanisms for contrast (Bonds, 1993) and luminance (Craik, 1940) then these should be investigated separately, to determine the actual root of the effects. This possibility will be investigated under more controlled circumstances, where contrast and absolute luminance level will be varied systematically. In the next experiment, the absolute luminance and contrast were varied. There were three levels of luminance; high, medium and low. Contrast was defined as a percentage of the maximum luminance, for example 100% contrast spans the range minimum to maximum luminance. 50% contrast would be a range of half this amount, etc. There were also two ranges of contrast about each mean luminance: a small contrast range (mean luminance +/- 12.5%) and a large contrast range (mean luminance +/- 25%). Additionally there was a contrast range spanning the entire possible luminance.
8.3 Experiment 2: CRT Monitor

This experiment was conducted to ascertain whether overall luminance level and the size of the contrast range affect discomfort and clarity judgements. This was using three classes of image: natural scenes, text and striped patterns. The data for Experiment 2 was collected by Caroline Foubister, who used the data and her own analysis as part of the work for her undergraduate degree qualification. The analysis, graphs, and discussion of these data in this thesis was done separately to the undergraduate project.

Method

Apparatus:
Stimuli were presented on a 19 inch Sony Trinitron Multiscan E400 monitor, calibrated using a Minolta LS-110 photometer. Screen size was 35.5cm x 26.4cm, resolution was 1024 x 768 pixels, and refresh rate was 60Hz. Participants viewed the display binocularly from a distance of 80cm. Luminance range was 1.3 to 108.2 cd/m².

Stimuli:
Stimuli were created and presented using Octave (Copyright J. W. Eaton and others, 2009, www.octave.org) and the Psychtoolbox ((Brainard, 1997; Pelli, 1997; Kleiner et al., 2007). Three types of stimuli were created: square-wave gratings, text and black and white natural images. The square-wave gratings were of spatial frequencies of 1,2,4,8,and 12 cycles/degree. Stimuli were presented at three base luminance levels; 27.4 cd/m², 54.1 cd/m² and 80.8 cd/m². Stimuli were varied in contrast by 25%, 50% and 100% of the total luminance range, around the mean value.

There were 35 gratings: three different luminance levels, three different contrast levels and five different spatial frequencies. There were three natural images taken from an internet site (www.storiesofplace.blogspot.com). The final stimulus was a custom made image of text (see Figure 8.3. The four images (three natural plus one text) were at three different luminance levels, the two extremes (27.4 and 80.8 cd/m²) had two contrast ranges (25% and 50%) and the middle luminance level (54.08 cd/m²) had 3 contrast ranges (25%, 50%, 100%), resulting in a total of 28 images.
8.3. EXPERIMENT 2: CRT MONITOR

The stimuli used in the contrast experiments. Three natural image pictures taken from www.storiesofplace.blogspot.com, and one custom-made Snellen-like letter chart.

Observers:
20 female and 4 male undergraduates with corrected-to-normal vision (mean age 21.6 years, range 21-23 years) participated in the experiment. Migraineurs and epilepsy sufferers were excluded from the study. All were naïve to the purposes of the experiment. All experiments were conducted according to the regulations of the University of St Andrews ethics committee (UTREC).

Procedure:
There were two conditions; in the first condition, observers were presented with square-wave gratings, in the second condition, observers were presented with photographs and text images. Stimuli were presented for 10s, after which time a mid-grey background was presented displaying the message ‘Rate picture COMFORT using keys 1-7, where 1 = very uncomfortable and 7 = very comfortable’. Observers responded using the number keys. After this response, another mid-grey screen was presented, this time requesting the observer ‘Rate picture CLARITY using keys 1-7, where 1 = not at all clear and 7 = very clear’. Participants again responded using the arrow keys. After these ratings were made, the next stimulus was presented for 10s.
8.3. EXPERIMENT 2: CRT MONITOR

Results

No participant’s data were excluded from analysis. Data relating to the 100% contrast range could only be collected around a single base luminance level of 54.1 cd/m$^2$. Therefore this condition was not included in statistical analysis of the effects of luminance level at constant contrast range. However, this condition was used in the analysis of the effects of different levels of contrast (25%, 50%, 100%) at the single luminance level of 54.1 cd/m$^2$. Repeated measures ANOVAs were performed: initially, a three-way luminance-by-contrast-by-frequency ANOVA to see if there were any effects of the experimental manipulations on either clarity, or comfort judgements.

Striped Patterns

The top row of Figure 8.4 shows the effect of spatial frequency of stripes on clarity judgements for each of the contrast ranges. Each individual plot is at constant luminance: left to right is lowest to highest mean luminance. The luminance-by-contrast-by-frequency ANOVA revealed a significant main effect of frequency ($F(4,92) = 34.88$, $p < 0.001$) and a significant luminance-by-contrast interaction ($F(2.0,45.1) = 12.60$, $p < 0.001$) on clarity judgements. From the top row of Figure 8.4 it can be seen that in all three plots (luminance levels), increasing spatial frequency decreases clarity judgements at all three contrast ranges. The interaction shows that there is a different direction of contrast effects at low luminance compared to high luminance. In Figure 8.4A and 8.4B it can be seen that lowest contrast range is judged to be clearer. However but this is reversed in Figure 8.4C, where higher contrast range is clearer.

The lower row of Figure 8.4 shows the effect of spatial frequency of stripes on comfort judgements for each of the contrast ranges. Again each individual plot is at constant luminance level. The luminance-by-contrast-by-frequency ANOVA revealed significant main effects of luminance ($F(1.6,36.8) = 30.84$, $p < 0.001$), contrast ($F(1,23) = 26.87$, $p < 0.001$) and frequency ($F(2.2,49.8) = 14.40$, $p < 0.001$). Comfort increased with increasing luminance, decreased with increasing contrast, and decreased with increasing spatial frequency. The interaction between luminance and contrast was significant ($F(2.0,46.0) = 19.57$, $p < 0.001$). As can be seen on the lower row of Figure 8.4D, 8.4E and
8.3. EXPERIMENT 2: CRT MONITOR

Figure 8.4: Clarity and comfort ratings of striped patterns. Top row shows clarity, lower row shows comfort ratings for gratings of varying spatial frequency. Contrast ranges are marked by the different lines. Each separate plot shows a constant luminance level. Error bars show ±1 standard error.

8.4F, the effects of luminance were greater for high contrast stimuli, and the effects of contrast were greatest at low luminance levels. This effect of contrast range is not apparent for the highest overall luminance level Figure 8.4F. The interaction between luminance and frequency was also significant (F(5.6, 129.8) = 2.73, p < 0.05). Spatial frequency tuning was clear at all luminance levels, but was more evident at low luminance levels (compare Figure 8.4D and 8.4E to 8.4F).

Natural Images
Comfort ratings were averaged across all three natural image pictures. The left hand side of Figure 8.5 shows mean clarity judgements against contrast range for the three levels of luminance. Results of a luminance-by-contrast repeated measures ANOVA showed significant effects of luminance (F(1.8, 42.0) = 59.33, p < 0.001) and contrast (F(1, 23) = 39.25, p < 0.001) on clarity judgements. There was also a significant interaction effect between luminance and contrast on clarity judgements (F(1.9, 42.5) = 7.14, p < 0.05). From the left hand side of Figure 8.5 it can be seen that clarity increases with contrast, and that, for
8.3. EXPERIMENT 2: CRT MONITOR

Figure 8.5: Clarity and comfort ratings for natural images.

Left plot shows clarity ratings, right plot shows comfort ratings for natural images at different contrast levels. Each line is a constant luminance level. Error bars show ± 1 standard error.

At a given level of contrast, low luminance images are judged to be clearer than others.

The right hand side of Figure 8.5 shows mean comfort judgements against contrast range for the three levels of luminance. Results of a luminance-by-contrast ANOVA showed there to be significant main effects of luminance ($F(1.8,41.3) = 7.40, p < 0.01$) and contrast ($F(1,23) = 12.34, p < 0.01$) on comfort judgements of natural images. There was also a significant luminance-by-contrast interaction effect on comfort judgements ($F(1.8,42.1) = 5.86, p < 0.01$). From the right side of Figure 8.5 it can be seen that there is an increase in comfort with contrast, but images of higher overall luminance tend to be judged as less comfortable than lower luminance images.
8.3. EXPERIMENT 2: CRT MONITOR

Figure 8.6: Clarity and comfort ratings for text stimuli.

Left plot shows clarity ratings for increasing contrast ranges for three luminance levels. Right plot shows comfort ratings for the increasing contrast range for three different luminance levels. Error bars show ± 1 standard error.

Text

The left side of Figure 8.6 shows clarity judgements for text stimuli. Results of a luminance-by-contrast repeated measures ANOVA showed there to be significant main effects of luminance (F(1.8,42.3) = 38.18, p < 0.001) and contrast (F(1.23) = 12.83, p < 0.01) on clarity judgements of the Snellen-chart-like image. From Figure 8.6, it can be seen that clarity ratings increase with contrast. Like the natural image stimuli, clarity ratings decrease with increasing overall luminance level.

The right side of Figure 8.6 shows the comfort ratings for text stimuli. Results of a luminance-by-contrast ANOVA show there to be a significant main effect of luminance only (F(1.7,39.7) = 4.10, p < 0.05). From Figure 8.6 it can be seen that the higher the overall luminance, the lower the comfort rating.
8.3. EXPERIMENT 2: CRT MONITOR

Discussion

Increasing spatial frequency of striped gratings reduced clarity judgements. This could be partially accounted for in terms of contrast sensitivity. Observers are around 10 times less sensitive to the higher spatial frequencies tested (12 cycles/degree) than to the lower ones in the range used (around 3 cycles/degree), which could be accounted for by contrast sensitivity (Campbell and Robson, 1968). However, this is only the case for the higher spatial frequencies. The lower ones in the range tested should have been less clear than the 3 cycles/degree stimulus, if contrast sensitivity was the cause. This was not the case, therefore contrast sensitivity cannot account for the complete effect. However, it is possible that there is a simple cause - better acuity may be needed to see the high spatial frequency stripes compared to the low spatial frequency stripes.

The higher contrast stripes were also judged as being clearer, when the overall luminance level was high. It is possible that contrast gain control can account for the results at low and medium luminance levels. Contrast gain control equalises the differences between sensitivity to higher and lower contrast ranges, thus they appear equally clear. However, the results do not entirely support this explanation, as this was not the case at the high mean luminance level: higher contrast ranges are clearer than low contrast ranges.

Comfort judgements of striped gratings were also affected by spatial frequency. This time there was a U-shaped tuning effect. This might again be explicable in terms of contrast sensitivity - observers are more sensitive to the midrange stimuli (Campbell and Robson, 1968). The peak of this function is expected to produce a larger cortical response than the other spatial frequencies. Additionally, higher contrasts should produce higher overall peaks, thus being more uncomfortable overall than lower contrast ranges. Again, this is consistent with the idea of increased cortical responses resulting in discomfort (Juricevic et al., 2010). However, again the effects are not the same for the highest mean luminance level, where there is a flattening of this tuning. One possible reason for the flattening at higher luminance levels is that these are at ceiling. This is not likely as the scale went up to ten, and these average around six. An alternative explanation is that spatial frequency tuning flattens with stimuli presented at higher contrast (suprathreshold) levels (Georgeson and Sullivan, 1975), which would flatten the discomfort across the spatial frequencies. However, as all were suprathreshold gratings, then this should not discriminate between the low and medium, and the highest mean luminance levels.
A different pattern of results was found for the natural image stimuli: both comfort and clarity judgements increased with contrast. One possible explanation for this is eyestrain from inefficient ocular motor responses. Sufficient contrast gradients are needed to drive ocular motor responses, for example, the accommodation response (Day et al., 2009a). Thus this could result in lower comfort ratings at lower contrast ranges. There was overall less comfort from the higher luminance stimuli. It is unclear why this is the case. One possibility is that luminance gain control processes reduce sensitivity overall, resulting in less clarity. A reduction in clarity might cause the ocular motor system to strain to maximise the contrast information needed to drive the response, thus causing discomfort.

The text stimulus showed a similar pattern as the natural images for clarity: a larger contrast range resulted in higher clarity judgements. However, while contrast ratings were higher at 50% contrast than at 25% contrast, no additional benefit was found for higher contrast levels. Again, higher overall luminance resulted in lower clarity and comfort, suggesting that the visual system might be straining to drive ocular motor responses from these stimuli, thus reduced clarity leads to reduced comfort. The effect of higher mean luminance reducing the clarity and comfort is again difficult to account for.

It is possible that hyperexcitation could account for this effect of reduced comfort at higher luminance levels. Higher contrast levels would cause larger cortical responses. Larger cortical responses have been suggested to be uncomfortable (Juricevic et al., 2010). However, hyperexcitation does not account for the effect of increasing luminance on clarity. Hyperexcitation might be caused by excessive noise. If there were noise in the visual system that increased with stimulus intensity, this might be expected to increase at higher luminance levels. This increase in noise might render the stimulus harder to detect. However, it is unclear why an increase in noise would not be accompanied by a corresponding increase in signal, therefore the signal-to-noise ratio and thus ability to detect the signal should remain constant.

There were effects of luminance on comfort and clarity using the range of possible luminances available on the CRT display. Additionally, the effects of contrast and spatial frequency appeared to be different at high mean luminance levels compared to lower mean luminance levels. Increasingly, high luminance displays are being developed, particularly for use in daylight conditions. A high-bright display makes it possible for presentation of stimuli at higher lu-
8.4 Experiment 3: High-Bright Display

Method

Apparatus:
In a third experiment, stimuli were presented on a 15 inch SLO 1568 LED backlit LCD monitor. The screen size was 30.4 x 22.8cm, and the screen resolution was 1024 x 768 pixels. Refresh rate was 60Hz. Luminance range was 4.07 to 1424.33 cd/m$^2$. The stimuli were viewed at a distance of 1m, which was measured using a piece of string to position the observers head at the appropriate distance.

Stimuli:
Stimuli were the same as in Experiment 2. Luminance range was 4.07 to 1424.3 cd/m$^2$. Three mean luminance levels were used; the high level varied around a midpoint of 1068.2, the medium luminance level around 712.2, and the low luminance level around 356.1 cd/m$^2$. There were two contrast ranges around each midpoint: 25% and 50%. The largest contrast range of 100%, which spanned the entire possible luminance range, was used for the medium luminance level only.

Observers:
24 student observers with corrected-to-normal vision took part in the study. Migraine and epilepsy sufferers were excluded from participating. Age range was 18 to 30.

Procedure:
Observers completed the same procedure as Experiment 2.

Results

Striped Patterns The upper row of Figure 8.7 shows clarity ratings against spatial frequency for each contrast range. Each plot shows results for a separate mean luminance level. Results of a luminance-by-contrast-by-spatial frequency repeated measures ANOVA showed there to be significant main effects of lumi-
8.4. EXPERIMENT 3: HIGH-BRIGHT DISPLAY

Figure 8.7: Clarity and comfort judgements of striped stimuli

Top: clarity judgements against increasing spatial frequency for each contrast range. Each plot is a separate luminance level. Lower: comfort judgements against increasing spatial frequency for each contrast range. Each plot is a constant luminance level. Errorbars show ± 1 standard error.

Clarity and comfort judgements of the striped gratings. There was also a luminance-contrast interaction effect (F(2.0,41.3) = 18.96, p < 0.001). From Figure 8.7 it can be seen clarity ratings decrease with spatial frequency. There is an effect of contrast, which can be particularly seen in Figure 8.7B: the lowest contrast range is judged to be less clear than the other two. Comparing across Figures 8.7A, B and C, it can be seen that the high and mid luminance levels appear to have higher clarity ratings than the low mean luminance level. The interaction is that the contrast effect can be seen for the medium luminance level only. There is little difference between the lines for each contrast level in the other two figures.

The lower row of Figure 8.7 shows comfort judgements against spatial frequency for each of the three contrast ranges. Each separate plot is a constant luminance level (lowest to highest, D to F). Results of a luminance-by-
8.4. EXPERIMENT 3: HIGH-BRIGHT DISPLAY

contrast-by-spatial frequency repeated measures ANOVA showed there to be significant main effects of contrast \( (F(1,21) = 20.25, p < 0.001) \) and spatial frequency \( (F(3.0,63.6) = 37.46, p < 0.001) \) on comfort judgements of striped patterns. There was also a significant contrast-by-spatial frequency interaction effect \( (F(3.3,68.6) = 6.08, p < 0.01) \), and a significant three-way interaction between luminance, contrast and spatial frequency \( (F(5.2,109.6) = 3.54, p < 0.01) \). The lower row of Figure 8.7 shows there to be a tuning effect of spatial frequency on comfort judgements. There is also an effect of contrast range, the lowest contrast range tends to be judged as being more comfortable across all three overall luminance levels.

Natural Images

The left side of Figure 8.8 shows mean clarity rating against contrast range for three luminance levels. Results of a luminance-by-contrast repeated measures ANOVA showed there to be significant main effects of luminance \( (F(1.6,34.4) = 7.64, p < 0.01) \) and contrast \( (F(1,21) = 23.54, p < 0.001) \) on clarity judgements of natural images. From Figure 8.8 it can be seen that clarity judgements increase with contrast. The medium luminance level is judged to be clearer than the other two.

The right side of Figure 8.8 shows comfort ratings against contrast range for the three luminance levels. Results of a luminance-by-contrast ANOVA showed there to be a significant main effect of contrast only \( (F(1,21) = 16.86, p < 0.01) \) on comfort judgements of natural images. From Figure 8.8 it can be seen that this is a tuning effect of contrast. There appears to be an increase in comfort with contrast that plateaus at 50% contrast range.

Text Stimuli

The left side of Figure 8.9 shows clarity judgements against increasing contrast range for three luminance levels. Results of a luminance-by-contrast repeated measures ANOVA show there to be a significant main effect of contrast only \( (F(1,21) = 30.28, p < 0.001) \) on clarity rating with increasing contrast which appears to plateau at 50% contrast.

The right side of Figure 8.9 shows comfort ratings against increasing contrast range for the text stimulus. Results of a luminance-by-contrast repeated measures ANOVA showed there to be a significant main effect of contrast only \( (F(1,21) = 13.48, p < 0.01) \) on comfort judgements of the text stimulus. From
8.4. EXPERIMENT 3: HIGH-BRIGHT DISPLAY

Figure 8.8: Clarity and comfort ratings of natural images.

Left: clarity judgements against increasing contrast range for three luminance levels for natural images. Right: comfort ratings against increasing contrast range. Errorbars show ± 1 standard error.
8.4. EXPERIMENT 3: HIGH-BRIGHT DISPLAY

Figure 8.9: Clarity and comfort ratings of text.

Left: clarity judgements against increasing contrast range for three luminance levels for the text stimulus. Right: comfort ratings against increasing contrast range. Errorbars show ± 1 standard error.
8.5. GENERAL DISCUSSION

Figure 8.7 it can be seen that this follows a similar pattern to the clarity judgements: comfort increases with contrast up to 50% contrast.

Discussion

The results with the high-bright monitor were similar to those of the ordinary CRT. The striped stimuli showed a decrease in clarity with increasing spatial frequency. The comfort judgements showed a tuning effect of contrast. Low contrast stimuli were less clear but also more comfortable.

Natural images also showed a similar pattern with the high-bright display as for the CRT. Clarity and comfort judgements once again increased with contrast. Comfort judgements appeared to plateau at 50% contrast, suggesting that at a certain level of contrast there is no extra benefit. This could be because the visual system has sufficient contrast information to inform the response. Text stimuli show a similar pattern to natural images: there is an increase in both comfort and clarity judgements up to 50% contrast, when there is no additional benefit.

The medium mean luminance level tended to be the clearest and rated most comfortable for the photographs of natural scenes. There was little difference in ratings between the luminance levels for text patterns. It is unclear why this should be the case. To speculate, higher luminance levels might have reduced sensitivity due to luminance gain control. Low luminance levels might be difficult to discriminate as signal is low. The difference between the photographs and the text might be because the black and white of the text patterns is intrinsically a high contrast stimulus, and easier to discern detail. Therefore, the medium luminance level might be the clearest and therefore most comfortable for the photographs, but this luminance effect is not seen for the intrinsically higher contrast text stimuli. Further work is needed to disambiguate the cause of these results.

8.5 General Discussion

Natural images and text patterns showed similar patterns, suggesting that in this case the text patterns are behaving like natural images, rather than periodic stimuli. This could be because the stimulus was a letter chart, and lacked the global configurations known to affect discomfort (Conlon and Humphreys, 2001). The letter chart was used as this was the same stimulus as used by Day et al. (2010), who have already conducted investigations into these displays.
The study therefore used the same stimuli to facilitate comparisons. It could be the case that due to the similarity of the clarity and contrast judgements that there is a common factor affecting them both. One possibility is that this is an ocular motor response: sufficient contrast might be needed to effectively drive the ocular motor response (e.g. Day et al., 2009a) and thus avoid eyestrain. The argument for ocular motor responses causing discomfort has some evidence to support it (e.g Day et al., 2009a; Chase et al., 2009; Tosha et al., 2009). Additionally the accommodation response is to some extent under voluntary control (Francis, Jiang, Owens, and Tyrrell, 2003; Taylor, Charman, O’Donnell, and Radhakrishnan, 2009) therefore it is possible that the additional effort of straining to see a stimulus contributes to the discomfort.

ocular motor response difficulty is only one possible explanation. Alternative arguments could be made to account for the findings, such as a low signal-to-noise ratio. A low signal-to-noise ratio might arguably transmit less useful information (Levy and Baxter, 1999). It is arguable that as a goal of the visual system is to extract information from images then this is not likely to be a comfortable stimulus to look at. These stimuli might also require further processing to extract the information, which could be metabolically costly. It might be expected that low signal-to-noise ratios would be uncomfortable stimuli, however, noisy stimuli are commonly used in psychophysical tasks for lengthy time periods (e.g. Neri, 2009). If low signal-to-noise levels were intrinsically uncomfortable, it would be expected that there will be more discomfort with poorer performance. However there is no evidence that stimuli with near-threshold performance on lengthy psychophysical tasks are considered uncomfortable.

Striped gratings did not show the same effects as natural images and text stimuli. Instead, stripes showed differential effects with contrast and clarity. Although increasing contrast increased clarity, this did not necessarily increase comfort. This could be explained in terms of contrast sensitivity (Campbell and Robson, 1968). The increased sensitivity to midrange spatial frequencies, such as 2-5 cycles/degree, could result in a larger cortical response to these spatial frequencies. This might lead to discomfort from excessive cortical responses (Juricevic et al., 2010).

There was little difference in discomfort reported from the two displays, even though one was twice the luminance of the other. This is likely to be due to effective gain control mechanisms. The human visual system can adapt to effectively process much higher luminances than that of the display (e.g. direct
8.5. GENERAL DISCUSSION

sunlight is several thousand cd/m²). Therefore it is perhaps unsurprising that
the high-bright display did not cause much discomfort in itself. Instead, far
more of a cause for concern is not the luminance level itself, but the range of
contrast possible with this type of display. Whilst natural images and text are
not themselves intrinsically uncomfortable stimuli, increasing the contrast is
beneficial as it increases clarity and comfort. However, over a certain level this
is not of any extra benefit (see Experiment 3).

However, the contrast of intrinsically uncomfortable stimuli should be moni-
tored. Although clarity increases with contrast for the striped stimuli, comfort
does not follow the same pattern. Therefore, caution should be excised before
presenting intrinsically uncomfortable stimuli at large contrast ranges.
CHAPTER 9

A WIDER LOOK AT VISUAL DISCOMFORT

9.1 Introduction

The results of Chapter 7 showed an association of discomfort group as defined by the VDS (Conlon et al., 1999) and reading ability - those who scored highly were both slower and made more mistakes at a reading task. This suggests that there is an association between susceptibility to visual discomfort and task performance. However, the results of Chapter 6 showed little effect of theoretically uncomfortable stimuli on performance on a search task. One possible explanation for this is that the stimuli themselves simply were not uncomfortable enough to affect performance in the wider population. Another explanation is that the participants were not susceptible enough to visual discomfort: it is possible that the sample of students might be more robust to discomfort, and that another sample drawn from the general population might be more sensitive to the stimulus manipulations. This is a possibility as there were very few observers reporting high levels of discomfort according to the Conlon et al. (1999) VDS in Chapter 7.

This low proportion of visual discomfort sufferers is not what was reported by Conlon et al. (1999). In their population of student nurses and their parents, it was shown that around 48% of participants reported moderate to high visual discomfort levels. Other studies by this research group report similar proportions of high visual discomfort sufferers in undergraduate samples: (see Conlon et al., 1998; Conlon and Hine, 2000; Conlon et al., 2001; Conlon and
9.1. INTRODUCTION

Humphreys, 2001). Other researchers also report similar-sized groups of high scoring individuals as classified by the VDS (e.g. Chase et al., 2009; Tosha et al., 2009).

It is possible that the lack of impact on performance on the search tasks shown in Chapter 6 could be because the sample of individuals was relatively more robust to visual discomfort. This could have been because the university environment, which is heavily biased towards reading, selects those who enjoy reading. It is logical that high visual discomfort sufferers might not enjoy reading, if they find it uncomfortable. Therefore high visual discomfort sufferers might be less inclined towards a university environment, as this tends to lay emphasis on reading. However, this seems an unlikely possibility, as the (Conlon et al., 1999; Chase et al., 2009; Tosha et al., 2009) findings all relate to students. It is possible that a larger sample from a wider population could be expected to include more high scorers on this scale. Therefore the current study was conducted to assess the susceptibility to discomfort in the general population, using a method designed to enable a greater variety of participants.

Types of Discomfort

There are several different aims of the survey. One is to investigate the possibility of there being different causes of discomfort that manifest themselves in different populations. For example, it has been suggested that myopic individuals and those over 40 are more tolerant of blur than emmetropic individuals (Wang and Cuiffreda, 2006). If problems with accommodation are a possible cause of discomfort (see Chapter 4), then these groups might report less discomfort, specifically from manipulations of stimulus blur. Conversely, those with very good optics might be less tolerant of blur than myopic individuals (Wang et al., 2006). Therefore individuals with normally good optics might be more sensitive than others to manipulations increasing image blur.

In an analogous way, it could be that individuals prone to large cortical responses might be more susceptible to discomfort from gratings. In the extreme, it has been suggested that those suffering migraine might have an overresponsive cortex (e.g. Aurora and Wilkinson, 2007), and therefore have more extreme reactions to striped gratings than shown by non-migraineurs. It is also possible that there is a spectrum of individuals with different levels of responsivity of the visual cortex. This might be shown by individuals being more susceptible to headache than others, whilst not necessarily suffering from migraine.
9.1. INTRODUCTION

This possibility might manifest itself in a wider distribution of individuals and might not have been seen in the limited samples of students from the previous chapters.

Alternative Measurements of Discomfort

The Conlon et al. (1999) scale is not the sole measure of visual discomfort. Another measure that is used is the Pattern Glare Test (Wilkins and Evans, 2001). This test has been shown to predict effectiveness of coloured visual overlays in improving reading speed (Hollis and Allen, 2006). The Pattern Glare Test consists of three striped gratings of differing spatial frequency, one of which is 3 cycles/degree. Observers answer questions regarding how many illusions, the types of illusions, and the discomfort induced by each of the gratings. The total score for the 3 cycles/degree grating is subtracted from the higher spatial frequency grating, to normalise the responses for observer response tendencies, such as differences in criterion. The Pattern Glare Test has been validated in a sample drawn from the general population (Evans and Stevenson, 2008). It would be useful to see if these two scales correlate. It is expected that they should, unless they were assessing different aspects of visual discomfort. It could be that, as one is mainly concerned with asking questions about reading, and the other requires observers to judge striped gratings, that these are actually two different aspects of visual discomfort. Thus scores might not correlate. It has yet to be tested whether or not they do correlate.

Predictability of Discomfort Ratings

A related aim of this survey is to investigate whether the reported discomfort from the kinds of stimuli used in the previous experimental chapters matches the susceptibility to discomfort according to the two scales. In other words, the extent to which standardised test scores reflect the ratings given on viewing theoretically uncomfortable stimuli. It could be the case that the manipulations increasing blur might separate from the manipulations proposed to affect cortical responsiveness. If these are two separate causes, it expected that these manipulations will group together in the discomfort judgements of individuals.

Comparing Across Stimulus Class

A final aim is to measure whether all classes of stimuli are equally uncomfortable. Previous research has focused on manipulations within a single type of stimulus, for example, striped patterns (Wilkins et al., 1984) or filtered noise (Fernandez and Wilkins, 2008; Juricevic et al., 2010). However, it is possible
that one class is overall more uncomfortable than another. For example, it might be the case that striped patterns, which have a history of being problematic for clinical populations (e.g. Radhakrishnan et al., 2005), might be more uncomfortable than the eyestrain caused by blurred stimuli, if this is indeed the case. Of course, this does not depend entirely on these theoretical explanations being correct - it could be the case that one class of stimulus is more uncomfortable than another for a variety of reasons. It would be useful to know if this is the case, as it could assist in comparing discomfort judgements across previous studies.

Aims

There are four aims of this study. The first is to measure discomfort using standardised tests, and see if the results of these correlate. The second is to see if the results of other experimental chapters, such as spatial frequency tuning (see Chapter 2) and blur effects (see Chapter 4) can be replicated in a wider population, outside a laboratory setting. The third is to compare discomfort judgements from these different potential types of discomfort. For example, it might be the case that stripes are more uncomfortable on the whole than blurred images. The final aim of this study is to investigate the predictive power of the standardised tests, and to see if age has any effect on discomfort. It might be expected that older people do not report so much discomfort from these stimuli. This could be because older adults have poorer optics. The accommodative response declines substantially after the age of 40 (Ostrin and Glasser, 2004). Therefore if it is indeed the case that accommodation provides a source of discomfort, then this should be reduced in older adults. This is only one possibility, another is reduced contrast sensitivity in older adults (Ross, Clarke, and Bron, 1985). Lowered apparent contrast might result in reduced discomfort judgements (see Chapter 8).

Questions unrelated to discomfort will also be used, so it can be seen if participants are simply responding to the image manipulation.

9.2 Method

Apparatus:
Stimuli were presented in an online survey using ‘Survey Monkey’ (http://www.surveymonkey.com/s/AbstractPictures, last accessed: October 2012, SurveyMonkey.com LLC, Palo Alto, California, USA).
9.2. METHOD

Observers:
Observers were recruited chiefly by word-of-mouth, and poster advertisement. As there is a large student population who were likely to receive this advertising, efforts were made to recruit additional age groups. In order to obtain older adults in the sample, some older groups were sent targeted email advertisements. This included participants taking part in a psychology evening degree, and other older adults were recruited as they had consented to be contacted through participation in previous research at the university. There was no reward incentive for participation. The survey was approved by the University ethics committee (UTREC). For ethical reasons, no participants under the age of 18 were permitted. Sufferers of epilepsy were also excluded from the study. Participants were free to omit any answers they wished. In total, 151 participants began the study. However, complete responses for the entire questionnaire (no omissions at all) were received from 87 participants. Mean age of respondents with complete answers was 27, SD = 9.43.

Stimuli:
The online questionnaire consisted of multiple components. The first component was the Conlon et al. (1999) Visual Discomfort Scale (VDS). This is a 23-item scale where respondents answer each question on a four-point scale: 0 for never, 1 for occasionally, 2 for often, 3 for always. The percentage of the total possible score is taken as the measure of discomfort. The second component was a measure of pattern sensitivity called the Wilkins and Evans (2001) Pattern Glare Test. Observers complete a series of short questions on three circular, horizontal square-wave striped patterns of differing spatial frequency. The 3 cycles/degree stimulus is either considered alone, or the scores from these are normalised by subtracting the 12 cycles/degree stimulus score. Stimuli were taken from previous experimental chapters. They were categorised into two main groups: manipulations of spatial frequency and manipulations of blur.

Spatial frequency manipulations consisted of ‘bump’ stimuli (see Chapter 2), striped gratings and harmonic manipulated stimuli (see Chapter 4). The bump stimuli were of spatial frequencies of 0.4, 7.5, 15.0, and 30.0 cycles/image (see Figure 9.1). The striped patterns were square-wave stripes of 12, 24, 48 and 96 cycles/image. These consisted of a five-component complex waveform. See Figure 9.2 for examples.
9.2. METHOD

Figure 9.1: Bump stimuli for the online survey
Bump stimuli of varying spatial frequency.

Figure 9.2: Stripe stimuli.
Stripe stimuli of varying spatial frequency.

Blur manipulated stimuli consisted of harmonic manipulated stimuli as used in Chapter 4, and defocus blurred natural images and filtered noise. The harmonic manipulated stimuli were radial waveforms with luminance profiles that were manipulated by removing harmonics: there was a fundamental-frequency-only sine waveform, and then stimuli with 1 harmonic, 2 harmonics, and three harmonics added. As in Chapter 4, it was expected that manipulations removing high spatial frequency content would increase discomfort judgements. Four of the phase-manipulated stimuli were also as Chapter 4. These were a five-component waveform with the phase of the first harmonic shifted by either 45, 90, 135, 180 degrees.

The natural images and $1/f$ filtered noise were blurred with 1.5, 2.5, 3.5 and 4.5 dioptres of defocus (see Equation 9.1). Defocus was diffraction-limited based on the equation of Levi and Austing (1968) and Levi (1974). This is an equation to calculate the change in contrast for a given level of defocus at a specified distance and spatial frequency. The reduction in contrast ($T$) is calculated for a diffraction-limited system. This is without any additional complications such as aberrations, which are present in biological systems. An additional simplification is the wavelength of light is specified, in this case.
greenish light (555nm wavelength). Thus the model ignores chromatic aber-
rations. The units of the original equation are normalised. Defocus is given
in Raleigh units, which are based on Raleigh’s resolution limit. This is the
minimum separation needed to resolve two points. Equations to convert the
Raleigh units into dioptres, and the spatial frequency into cycles/degree, are
given in Equations 9.3 and 9.4 respectively.

\[
T(\nu r, \delta) = \frac{4}{a\pi} \int (1 - x^2)^{0.5} \cos |a(x - \nu r)|dx \tag{9.1}
\]

Where

\[
a = 2\pi \nu r \delta \tag{9.2}
\]

Where \(\delta\) is defocus in units based on Raleigh tolerance, and \(\nu r\) is the nor-
malised spatial frequency. Equations for conversion are given below:

\[
\delta = \frac{d^2D}{2 \times 10^{-3}\lambda} \tag{9.3}
\]

\[
\nu r = \frac{\lambda f}{1.746 \times 10^4 d} \tag{9.4}
\]

Where \(D\) is the units for defocus in dioptres, \(\lambda\) is the wavelength of light
(assumed here to be 555nm), \(f\) is the spatial frequency and \(d\) is the pupil
diameter in mm.

It was expected that the more blurred the stimuli, the more discomfort would
be reported.
9.2. METHOD

Figure 9.3: Harmonic-manipulated radial waveform stimuli.

Left to right: radial waveforms with varying numbers of harmonics: 4-component, 3-component, 3-component, sine wave fundamental only.

Figure 9.4: Optically blurred natural images.

Natural images with increasing optical blur. Optical blurring was based on a diffraction-limited model. Left to right: 1.5D, 2.5D, 3.5D, 4.5D. 1 pixel = 1 arcmin in the original image.

Figure 9.5: Optically blurred filtered noise.

$1/f$ filtered noise with increasing optical blur. Optical blur was based on a diffraction-limited model. Left to right: 1.5D, 2.5D, 3.5D, 4.5D. 1 pixel = 1 arcmin in the original image.

Phase manipulated stimuli used in Chapter 4 were also employed. As there was no change in the spatial frequency content, it was expected that these stimuli should not show a difference in discomfort judgements.
9.3. RESULTS

Figure 9.6: Phase manipulated radial waveform stimuli.

Radial waveform stimuli with varying phase shift of first harmonic: Left to right: 45, 90, 135 and 180 degrees first harmonic phase shift.

It must be noted that although these were set to be of particular spatial frequencies, this was only the case when viewed at 40cm. Differences in resolution, display screen size, and viewing distance cannot be accounted for, and therefore the exact spatial frequencies and blur figures quoted are not accurate representations of those that would have been presented to the observers. This was beyond experimental control in the current set-up. Therefore only qualitative conclusions can be drawn.

Procedure:
Observers first completed the VDS, then the Pattern Glare Test. Observers then rated images for discomfort, how interesting they were, how much they liked them, and how magnified they appeared. It is expected that discomfort and liking might not correlate: experiments have shows that aesthetic judgements that might not necessarily be the opposite of discomfort (Juricevic et al., 2010). How magnified the object appeared was a question relating to blur perception. It was thought that if observers felt the image was highly magnified then they might expect a certain level of blur to be present in the image due to DOF effects (see Chapter 5). Interest was a catch question designed to assess whether individuals were simply responding to image manipulation. Finally, observers completed some questions on demographics and visual health: age, need for optical correction, enjoyment of reading, and vision needed for occupation. Questions can be seen below in the Results section.

9.3 Results

Demographics

98 people answered the questions regarding demographics and general vision. The results can be seen in Table 9.1. Around half the participants reported
9.3. RESULTS

Table 9.1: Results for the general vision questions.

<table>
<thead>
<tr>
<th>Question</th>
<th>Distance</th>
<th>Reading</th>
<th>All the time</th>
<th>Don't need</th>
</tr>
</thead>
<tbody>
<tr>
<td>Do you need glasses?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Do you have headaches?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Do you enjoy reading?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>What kind of job?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Question</th>
<th>Frequency</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Do you need glasses?</td>
<td>Distance</td>
<td>Reading</td>
<td>All the time</td>
<td>Don't need</td>
</tr>
<tr>
<td>Do you have headaches?</td>
<td>once a week</td>
<td>month</td>
<td>seldom</td>
<td></td>
</tr>
<tr>
<td>Do you enjoy reading?</td>
<td>enjoy</td>
<td>Work only</td>
<td>Avoid</td>
<td></td>
</tr>
<tr>
<td>What kind of job?</td>
<td>Manual</td>
<td>Office-based</td>
<td>Need fine vision</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Question</th>
<th>Frequency</th>
<th>Frequency</th>
<th>Frequency</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Do you need glasses?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Do you have headaches?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Do you enjoy reading?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>What kind of job?</td>
<td>Manual</td>
<td>Office-based</td>
<td>Need fine vision</td>
<td></td>
</tr>
</tbody>
</table>

Table 9.1: Results for the general vision questions.

There are significantly more people who wear glasses all the time as a correction ($\chi^2 = 15.39, N = 98, p < 0.01$). There are significant differences between the number of people suffering from headaches seldomly, and the number of those who report more frequent headaches ($\chi^2(2) = 83.41, N = 98, p < 0.01$). There are also significantly more people who like reading than who avoid reading or read only for work purposes ($\chi^2(2) = 118.91, N = 98, p < 0.01$). Significantly more people have office-based jobs than manual jobs or jobs re quiring fine vision ($\chi^2(2) = 107.53, N = 86, p < 0.01$).

Picture Ratings

Discomfort Within Each Stimulus Class

Discomfort ratings across each stimulus class were analysed. Figure 9.7A shows discomfort judgements for bump stimuli. Tuning effects similar to those shown in previous experiments can be seen: the midrange was judged to be most uncomfortable. Results of a Friedman’s test showed this tuning to be significant for the bump stimuli ($\chi^2(3) = 41.43, N = 115, p < 0.01$). Discomfort for the striped stimuli are plotted in Figure 9.7B. Again, there is tuning, with a drop in discomfort at the low spatial frequencies. Results of a Friedman’s test again showed this tuning to be significant for the striped patterns ($\chi^2(3) = 104.56, N = 100, p < 0.01$). Figure 9.7C shows no effect of phase on discomfort judgements ($\chi^2(3) = 6.55, N = 100, p = NS$).
9.3. Results

Figure 9.7: Discomfort judgements.

Discomfort judgements against stimulus manipulation for each stimulus class. Top row: A: bump stimuli of increasing spatial frequency, B: stripe stimuli of increasing spatial frequency, C: phase manipulation of increasing shift of first harmonic. Lower row: D: harmonic stimuli with decreasing number of harmonics, E: natural images with increasing defocus blur, and F: noise images with increasing defocus blur.

For the blurred stimuli there is more of a monotonic increase in discomfort with increasing blur manipulation (see Figure 9.7E). This is in agreement with previous work (see Chapter 4). Results of Friedman’s tests show there to be a significant effect of blurring through loss of harmonics ($\chi^2(3) = 41.49$, $N = 110$, $p < 0.01$), for natural images ($\chi^2(3) = 203.59$, $N = 98$, $p < 0.01$), and for filtered noise ($\chi^2(3) = 87.99$, $N = 104$, $p < 0.01$).

Interest for Each Stimulus Class

Figure 9.8A shows the effect of bump stimulus spatial frequency manipulations on interest ratings. Figure 9.8B shows the effect of stripe spatial frequency on interest ratings. Figure 9.8C shows the effect of increasing the phase shift of the first harmonic on interest ratings. Although relatively flat lines, the effect of bump spatial frequency on interest ratings was significant according to a Friedman’s test ($\chi^2(3) = 25.57$, $N = 115$, $p < 0.01$). The effect of phase
9.3. RESULTS

Figure 9.8: Interest judgements.

Interest judgements against against stimulus manipulation for each stimulus class. Top row: A: bump stimuli of increasing spatial frequency, B: stripe stimuli of increasing spatial frequency, C: phase manipulation of increasing shift of first harmonic. Lower row: D: harmonic stimuli with decreasing number of harmonics, E: natural images with increasing defocus blur, and F: noise images with increasing defocus blur.

manipulation on interest ratings was also significant ($\chi^2(3) = 19.76, N = 100, p < 0.01$). However, the effect of stripe spatial frequency on interest judgements was not significant ($\chi^2 = 1.96, N = 100, p = \text{NS}$).

Figure 9.8D, E, and F show the effect of increasing blur manipulation on interest ratings. Interest judgements increase with blurring. The effect of blurring on interest was significant according to a Friedman’s test for the harmonic manipulation ($\chi^2(3) = 88.92, N = 110, p < 0.01$), for the natural images ($\chi^2(3) = 154.20, N = 98, p < 0.01$), and for the filtered noise ($\chi^2(3) = 24.06, N = 104, p < 0.01$).

Liking judgements for Each Stimulus Class

Figure 9.9A shows the effect of bump stimulus spatial frequency manipulations on interest ratings. There is a tuning effect of bump spatial frequency on liking
9.3. RESULTS

Liking judgements against against stimulus manipulation for each stimulus class. Top row: A: bump stimuli of increasing spatial frequency, B: stripe stimuli of increasing spatial frequency, C: phase manipulation of increasing shift of first harmonic. Lower row: D: harmonic stimuli with decreasing number of harmonics, E: natural images with increasing defocus blur, and F: noise images with increasing defocus blur.

ratings. Results of a Friedman’s test show this is significant ($\chi^2(3) = 44.20$, $N = 115$, $p < 0.01$). Figure 9.9B shows the effect of stripe spatial frequency on interest ratings. Increasing spatial frequency of stripes results in increased liking ($\chi^2(3) = 23.07$, $N = 100$, $p < 0.01$). Figure 9.9C shows the effect of increasing the phase shift of the first harmonic on liking ratings. This too is significant according to the results of a Friedman’s test ($\chi^2(3) = 8.54$, $N = 100$, $p < 0.01$).

Figure 9.9D, E, and F show the effect of increasing blur manipulation on interest ratings. Liking judgements increase with blurring. The effect of blurring on liking was significant according to a Friedman’s test for the harmonic manipulation ($\chi^2(3) = 43.94$, $N = 110$, $p < 0.01$), for the natural images ($\chi^2(3) = 181.46$, $N = 98$, $p < 0.01$), and for the filtered noise ($\chi^2(3) = 66.26$, $N = 104$, $p < 0.01$).
9.3. RESULTS

Figure 9.10: Magnification judgements.

Magnification judgements against stimulus manipulation for each stimulus class. Top row: A: bump stimuli of increasing spatial frequency, B: stripe stimuli of increasing spatial frequency, C: phase manipulation of increasing shift of first harmonic. Lower row: D: harmonic stimuli with decreasing number of harmonics, E: natural images with increasing defocus blur, and F: noise images with increasing defocus blur.

Magnification judgements for Each Stimulus Class

Figure 9.10A shows the effect of bump stimulus spatial frequency manipulations on how magnified the stimulus was considered to be. Although apparently fairly flat, there is a significant effect of spatial frequency ($\chi^2(3) = 20.16$, $N = 115$, $p < 0.01$). Figure 9.10B shows the effect of stripe spatial frequency on magnification ratings. Magnification judgements decrease with increasing spatial frequency ($\chi^2(3) = 47.59$, $N = 100$, $p < 0.01$). Figure 9.10C shows the effect of increasing the phase shift of the first harmonic on ratings of how magnified the observer considered the stimulus to be, this is also a significant effect according to Friedman’s test ($\chi^2(3) = 21.02$, $N = 100$, $p < 0.01$).

Figure 9.10D, E, and F show the effect of increasing blur manipulation on interest ratings. The harmonic manipulated stimuli were considered to be

215
9.3. RESULTS

less magnified with loss of high spatial frequency information ($\chi^2(3) = 82.74$, $N = 110$, $p < 0.01$). How magnified the observer considered the defocussed natural and noise stimuli to be increased with blurring. The effect of blurring on magnification was significant according to a Friedman’s test for the natural images ($\chi^2(3) = 62.84$, $N = 98$, $p < 0.01$), and for the filtered noise ($\chi^2(3) = 139.50$, $N = 104$, $p < 0.01$). The effect is in different directions for the harmonic manipulations (see Figure 9.10D) compared to the defocused natural images and noise (see Figure 9.10E and F). It is unclear why this should be the case.

Comparing Across Stimulus Classes

In order to compare across stimulus classes, two different types of analysis will be done. Firstly the total rating of all four in a particular class of stimuli will be calculated. This will obtain a total score, for example discomfort rating, for all four stimuli within a particular class. Secondly, the most uncomfortable stimulus according to the results of the picture ratings (see above) will be taken. This will allow for comparisons without the potential dilution of other stimuli within the class that are perhaps not so uncomfortable. It could be the case that there is one particularly uncomfortable stimulus within a class, whilst the others are less uncomfortable. The total score would potentially reduce this effect if the others do not score so highly.

Discomfort judgements:
The total discomfort judgement across all four stimuli was calculated for each individual observer. The mean (across individuals) of these total discomfort judgements for each stimulus class can be seen in Figure 9.11A. This shows there to be clear differences between the stimulus categories in terms of total discomfort judgements. Stripe, harmonic and phase stimuli are judged as more uncomfortable than the bump, blurred natural images and blurred filtered noise. This effect of stimulus category on discomfort judgements is significant according to a Friedman’s repeated measures analysis of variance for non-parametric data ($\chi(5) = 209.15$, $N = 98$, $p < 0.01$). Post-hoc comparisons (Wilcoxon Signed Rank Test) showed there to be significant differences between stripe and harmonic stimuli ($T(98) = -3.51$, $p < 0.001$). There is no significant difference between noise and blurred natural images in terms of discomfort ratings ($T(98) = 0.24$, $p = NS$). Also there is no difference between noise and bump stimuli ($T(98) = 0.98$, $p = NS$). However, there is a significant difference between noise and phase stimuli ($T(98) = 4.81$, $p < 0.001$).
The most uncomfortable stimulus in each set according to the discomfort judgements (see previous section) was chosen. This was the 15 cycles/image bump stimulus, the 48 cycles/image striped grating, the 45° phase shifted stimulus, the sine wave, and the 4.5 dioptres blur added to the natural images and the filtered noise. Figure 9.11B shows the mean discomfort judgements for the most uncomfortable stimulus in each stimulus set. Again the striped pattern is the most uncomfortable stimulus class. However, this time the other stimuli are grouped much closer together. The effect of stimulus category on discomfort judgements of the most uncomfortable stimuli was significant ($\chi^2(5) = 115.38$, $N = 100$, $p < 0.01$). Post-hoc Wilcoxon Signed Rank tests showed there to be significant differences between stripes and phase stimuli for discomfort ($T(100) = -7.08$, $p < 0.001$). However, there is no significant difference between bump and noise stimuli ($T(100) = -0.61$, $p = NS$), although there is a difference between bump and phase stimuli in terms of discomfort judgements ($T(100) = 2.01$, $p < 0.05$).
9.3. RESULTS

![Graph](image)

Figure 9.12: Comparing interest across stimulus class

A: total interest judgements for all four stimuli in a class. B: interest judgements for the most uncomfortable stimulus in each class.

Interest judgements:

Total interest judgements can be seen in Figure 9.12A. There are clear differences between interest judgements of the different stimulus classes, significant according to a Friedman’s test ($\chi^2(5) = 271.43$, $N = 98$, $p < 0.01$). Bump and defocused noise stimuli are judged most interesting. This is followed by striped patterns and the harmonic stimuli. Results of post-hoc Wilcoxon Signed Ranks test showed that there is no significant difference between bump and noise stimuli ($T(98) = -1.20$, $p = NS$). However, bump stimuli are significantly more interesting than stripe stimuli ($T(98) = -2.80$, $p < 0.05$). Phase is the least interesting, with a significant difference between interest ratings for phase and blurred natural images ($T(98) = 3.75$, $p < 0.001$).

Figure 9.12B shows the interest judgements for the most uncomfortable stimulus in each class. Again, this shows maximum interest judgements for bump, filtered noise and stripe stimuli. However, this time the harmonic stimuli are not considered so interesting under this analysis. Effects are significant according to a Friedman’s test ($\chi^2(5) = 191.31$, $N = 98$, $p < 0.01$). Post-hoc
9.3. RESULTS

Figure 9.13: Comparing liking across stimulus class

A: total liking judgements for all four stimuli in a class. B: liking judgements for the most uncomfortable stimulus in each class.

tests (Wilcoxon Signed Ranks) showed there to be significantly higher interest ratings for bump stimuli compared to noise ($T(98) = -1.97$, $p < 0.05$). There was no significant difference between phase and harmonic manipulated stimuli ($T(100) = 0.57$, $p = NS$). However there was a significant difference between interest ratings for phase and blurred natural images ($T(100) = 2.36$, $p <0.05$).

Liking judgements:
Figure 9.13 shows the total liking judgements for each stimulus class. There are significant differences between the stimulus classes ($\chi(5) = 141.48$, $N = 98$, $p < 0.01$). The striped stimuli are the most liked stimulus. Post-hoc Wilcoxon Signed Ranks test show that blurred natural images are the least liked stimuli, there is a significant difference between blurred natural images and phase stimuli ($T(98) = 2.18$, $p < 0.05$). Stripe stimuli are significantly more liked than the bump stimuli ($T(98) = 2.42$, $p < 0.05$).

Figure 9.13B shows that the striped and bump stimuli are the most liked, even when analysing the most uncomfortable stimulus for each stimulus category.
Results of a Friedman’s test show this to be significant effect \((\chi(5) = 127.98, N = 97, p < 0.01)\) of stimulus category on liking. Post-hoc tests showed that the phase manipulated stimulus was the least liked: there was a significant difference between phase and the next lowest rating of blurred natural images \((T(96) = 2.97, p < 0.05)\). There was no significant difference between the top two for liking, bump and stripe stimuli \((T(96) = 0.07, p = \text{NS})\). There is no significant difference between noise and harmonic \((T(96) = -1.24, p = \text{NS})\). However, there was a significant difference between noise and blurred natural images \((T(96) = -2.50, p < 0.05)\).

**Magnification judgements:**
Figure 9.14A shows the total magnification judgements for each stimulus category. The bump stimuli are considered to be most magnified over all stimulus categories. Effect of stimulus category on magnification judgements is significant according to a Friedman’s test \((\chi(5) = 154.65, p < 0.01)\). Results of post-hoc Wilcoxon Signed Ranks test showed there to be no significant difference between blurred natural images and phase manipulated stimuli \((T(98) = 0.14, p = \text{NS})\) or between blurred natural images and striped patterns \((T(98) = 1.08, p = \text{NS})\). However, there was a significant difference between blurred natural images and noise \((T(98) = 2.25, p < 0.05)\).

Figure 9.14B shows the magnification judgements for the most uncomfortable stimulus of each stimulus class. Again, the bump stimuli are shown to be the most magnified. The most uncomfortable harmonic and filtered noise stimuli are considered to be more magnified that the stripes and the phase manipulated stimuli. The effect of category on magnification judgements are significant according to a Friedman’s test \((\chi(5) = 127.98, p < 0.01)\). Results of post-hoc Wilcoxon Signed Ranks test showed there to be no significant difference between stripe and phase stimuli \((T(97) = -2.11, p = \text{NS})\). However, there is a significant difference between phase and blurred natural images \((T(97) = 2.84, p < 0.05)\). There was a significant difference between bump and harmonic stimuli \((T(97) = -3.47, p < 0.05)\).

**Measures of Discomfort**

**Visual Discomfort Scale**
The VDS is a 23-item questionnaire. Participants respond by answering 0 (never) to 3 (always) for each item. The overall percentage score is then taken.
9.3. RESULTS

Figure 9.14: Comparing magnification judgements across stimulus class

A: total magnification judgements for all four stimuli in a class. B: magnification judgements for the most uncomfortable stimulus in each class.

Individuals scoring over 35% are classified as moderate discomfort sufferers, those over 70% are categorised as high discomfort sufferers.

119 respondents completed the VDS. The highest overall score was 43.5%, and the median score was 10.2%. 6% of the 119 respondents were classified as moderate discomfort sufferers. No individual scored highly enough to be classified as a high visual discomfort sufferer. By contrast, Conlon et al. (1999) reported 48% of 177 observers scoring highly enough to be classified as moderate or high discomfort sufferers.

Pattern Glare Test

The Pattern Glare Test (Wilkins and Evans, 2001) involves reporting the effects perceived while observing high contrast square-wave gratings. Effects include discomfort, pain, illusions of shape and colour, scintillation and movement. The number of effects reported from the 3 cycles/degree grating is taken as the measure. This is the raw score. This can be normalised for observer response bias by subtracting the score for the 12 cycles/degree grating. This is the normalised score. The raw score and the normalised score were compared. There was a significant, though small correlation between raw and normalised...
9.3. RESULTS

scores ($\rho(117) = 0.17$, $p < 0.05$).

There were 14% of respondents who obtained a high score (more than 2) on the normalised Pattern Glare Test, and there were 26% of respondents who obtained a high raw score. This is increased on the 5% observers reported by Evans and Stevenson (2008).

117 observers completed both the VDS and the Pattern Glare Test. The two measures correlate positively when the raw scores are considered ($\rho(117) = 0.20$, $p < 0.05$). However there is a negative correlation when the normalised scores are considered ($\rho(117) = 0.27$, $p < 0.01$).

**Susceptibility to Discomfort:**

One of the aims of this survey was to assess the predictive ability of the standardised tests of discomfort judgements. It was thought that if the VDS and the PS are measuring susceptibility to visual discomfort, it should be the case that there will be a correlation between the discomfort judgements for the pictures and the scores each individual attains on each standardised test. Also, it was predicted that the older population might report less discomfort than younger individuals. Thus there might be a negative correlation of discomfort judgements and age. The total discomfort rating of each of the four stimuli in each class was calculated. This was correlated against the VDS score and raw PS score for each individual. Also, the total discomfort was correlated against the age of the individual. The correlation coefficients are plotted in Figure 9.15.

Total discomfort ratings of the patterns correlated well against the VDS for some stimulus classes. Correlation coefficients significant at $p < 0.05$ are indicated with a star. Raw PS correlated fairly well for these patterns also. This same pattern can be seen in Figure 9.16. This is the same analysis, but this time instead of the total discomfort across the entire class of stimuli, Figure 9.16 plots the stimulus that should theoretically be the most uncomfortable (see previous section). It seems to be the case that pattern sensitivity, and VDS scores correlate with discomfort judgements of these images.

The interest judgements, which was intended to be the catch question, does not
9.3. RESULTS

Figure 9.15: Discomfort judgements correlated with VDS, raw PS score and age.

Correlation coefficients for Pearson’s correlations of total discomfort from a category and VDS, PS, and age. Significant (p < 0.05) coefficients marked with a star.

Figure 9.16: Discomfort judgements correlated with VDS, raw PS score and age.

Correlation coefficients of theoretically most uncomfortable stimulus in the category and VDS, PS and age. Significant correlation coefficients indicated with a star.
9.3. RESULTS

Figure 9.17: Interest judgements correlated against VDS, raw PS score and age.

Correlation coefficients of interest judgements and VDS, PS and age. Significant coefficients marked with a star.

show the same pattern of results as the discomfort judgements. How interesting the pattern is does not correlate well with either VDS or raw PS score. This can be seen in Figure 9.17. This could indicate that discomfort judgements are measuring something stable, as these three measurements correlate, whereas this is not the case for interest ratings.

However, the normalised version of the discomfort judgement shows a different pattern of results. Figure 9.18 shows the same correlation coefficients as before, this time with the normalised PS score, rather than the raw score. From this plot it can be seen that the pattern of results is no longer the same - pattern sensitivity score is now either non-significant or negatively correlated with the stimulus discomfort judgements. This is important as it suggests that discomfort judgements might be measuring the same thing, however this could be accounted for by observer response tendencies.

Another similar reversal of the previous results is seen for the interest judgements. Figure 9.19 shows the same correlation coefficients as before, this time for the normalised pattern sensitivity scores.
9.3. RESULTS

Figure 9.18: Discomfort judgements correlated against VDS, normalised PS score and age.

Top: Correlation coefficients of discomfort judgements and VDS, normalised PS and age for the total interest judgements for that stimulus category. Lower: Correlation coefficients of discomfort judgements and VDS, normalised PS and age for the theoretically most uncomfortable stimulus.
Figure 9.19: Interest judgements correlated against VDS, normalised PS score and age.

Top: Correlation coefficients of interest judgements and VDS, normalised PS and age for the total interest judgements for that stimulus category. Lower: Correlation coefficients of interest judgements and VDS, normalised PS and age for the theoretically most uncomfortable stimulus.
9.4 Discussion

Results of the survey showed tuning effects for discomfort judgements of the striped patterns. Most uncomfortable stimulus appeared to be the midrange of the set of stimuli. Increasing blur manipulations showed increased discomfort. These data show a similar pattern of results to those of the individual experimental chapters (see Chapters 2 and 4). However, the stimuli showing increased discomfort judgements were also those showing increased interest judgements. As this was the catch question, it was expected that the responses would show no particular pattern. As this was not the case, then caution must be used in interpreting the results further, as there is the danger that participants were responding simply to image manipulation in this case.

The total discomfort judgments for each stimulus class correlated with scores on the VDS, and on the raw scores on the PS. Scores on the VDS and raw scores on the PS scale did not correlate so much with interest judgements. This indicates that both measures are assessing something meaningful and similar, and are predictive of discomfort judgements of images. However, as the normalised scores on the PS did not correlate well with the VDS, or with the discomfort ratings of the stimuli, perhaps it is not actually measuring discomfort. The purpose of normalising the discomfort ratings is to account for observer response bias. For example, some observers might tend towards always reporting affirmatively, whilst others might always report more conservatively, despite their actual experiences being the same. Normalisation should reduce the impact of response bias on the data. As the effect does not remain after normalisation, this seems to indicate that all that is being measured is a tendency for observers to respond positively or negatively to stimulus manipulations.

Age was also predictive of discomfort judgements. However, contrary to expectations, age correlated positively with discomfort judgements. It is unclear why this should be the case.

The most uncomfortable stimulus class overall was the class consisting of the striped patterns. However, the striped patterns were also the most liked stimuli. The stimulus classes considered the most interesting tended to be the ones composed of artificial stimuli, rather than the blurred natural images.
9.4. DISCUSSION

Additionally, the stimuli appearing to be most magnified seemed to be the noise patterns. It is plausible that these stimuli were considered to be so as observers had no idea of scale, whereas with the natural images they did, for example. However, as the control questions showed effects, it is not possible to interpret the discomfort judgements as being valid.

Only very few observers reported high levels of discomfort according to the Conlon et al. (1999) VDS. There were no high visual discomfort scorers in the sample according to the classification criteria. This is curious that the current sample is apparently more robust to discomfort than those tested by previously by Conlon et al. (1999); Chase et al. (2009); Tosha et al. (2009). It could be the case that there are cultural differences in the interpretation and expression of discomfort.

The lack of high visual discomfort sufferers in the sample could go some way in explaining the lack of findings in Chapters 6 and 8. Repeating these experiments with a clinical population, such as migraine sufferers, might result in more extreme results. However, from these results it cannot be concluded that there is a large problem to the general population posed by uncomfortable stimuli.
CHAPTER 10

MIGRAINE

10.1 Introduction

The results of the survey of the general population in the previous chapter showed that very few individuals report large amounts of discomfort. However, more extreme responses might be expected for some clinical groups, particularly those associated with light sensitivity, such as migraine (e.g. Khalil, 1991) and epilepsy (e.g Wilkins et al., 1975; Radhakrishnan et al., 2005).

Migraine is a disorder that has strong associations with vision: some migraine sufferers report visual triggers of attacks, or visual abnormalities during attacks, and there are differences in visual performance compared to controls between attacks. In some migraine sufferers, the attack is heralded by a set of sensory disturbances, visual in many sufferers, known as the migraine aura (e.g. Sacks, 1992).

Additionally, there have been suggestions that the cause of migraine might involve a wavefront of excessive cortical activity that advances through the cortex, known as the cortical spreading depolarisation (CSD), which advances at the same rate across the cortex as the spread of the aura across the visual field. This is relevant to visual discomfort as it has been suggested that excessive activity resulting from particular visual stimuli might be one of the possible causes of increased discomfort judgements in non-clinical populations (Juricevic et al., 2010).

The purpose of this chapter is to discuss the potential theoretical reasons why migraine sufferers might show more extreme discomfort responses. Firstly,
10.2. Features of Migraine

Migraine is not an uncommon disorder, estimated to affect 15% of the population (Steiner, Scher, Stewart, Kolodner, Lieberman, and Lipton, 2003). Individuals reporting migraine tend to be predominantly women: 7.6% men surveyed suffered from migraine compared to 18.3% women (Steiner et al., 2003), although there were more women (approximately 63%) participating in this study overall than men (37%). Other research found there to be approximately equal numbers of male and female migraine sufferers with random (as opposed to self-selected) sampling (Yoon, Mueller, Hansen, Poitz, Slomke, Dommes, Diener, Katsarava, and Obermann, 2010).

Migraine is typically perceived as being a severe headache, however, migraine is not limited to this: the headache is typically accompanied by nausea, sensory disturbances, photophobia (particular sensitivity to light) and phonophobia (heightened sensitivity to noise). The International Headache Society diagnostic criteria for migraine (classification committee if the IHS, 1988) (Headache Classification Subcommittee of the International Headache Society, 2004) requires at least five attacks of lateralised, pulsating, headache lasting four or more hours, which disrupts daily activities, and is aggravated by physical activity. The headache is to be accompanied by at least one of the following: nausea, photophobia or phonophobia. In the case of some rare types of migraine (e.g. familial hemiplegic migraine) there is a temporary loss of control of movement during the attack.

Environmental Causes of Migraine

Certain stimuli in the environment can elicit migraine attacks. Migraine triggers are highly idiosyncratic, but there are some general trends; the patients of Sacks (1992) report triggers to include certain foods, stress levels, sleep deprivation, and also some visual stimuli. A large study of 1750 migraineurs conducted by Kelman (2007) found stress to be the most common environmental trigger of migraine; nearly 80% of migraineurs reported that stress was a factor in eliciting attacks. Other factors reported to be common triggers were
hormone levels, weather, too little or too much sleep, food, alcohol, fasting, exercise and, critically, light.

**Individual Variation in Migraine**

There are many manifestations of collections of symptoms. The duration and severity of symptoms is also highly variable between individuals, as is the time period between attacks. Sacks (1992) reports migraine patients whose migraines last for a few hours, and also less fortunate individuals who suffer for days. Sacks (1992) investigated many case studies of his patients who suffered from migraine, and these studies can give a much better idea of the incredible variety of manifestations of migraine. This variety means there is often some uncertainty in diagnosis. Migraine is a very heterogeneous disorder, with many subtypes. Additionally, related disorders, including cluster headache (bouts of frequent severe headaches), tension headache, and familial hemiplegic migraine, further complicate the issue of diagnosis. An unrelated disorder, occipital epilepsy, is commonly misdiagnosed as migraine (Panayiotopolos, 1999).

For some sufferers, the migraine is preceded by sensory disturbances, the ‘migraine aura’, that develop over a period of 5-20 minutes and last for an hour. These sensory disturbances form the main classification of migraineur subtype: migraine with (MA) and migraine without (MO) aura. These sensory disturbances can be across any modality, but the most commonly reported type of sensory disturbance is visual, therefore this is of particular interest to the current work.

**10.3 Migraine Aura**

According to a longitudinal study, 39% of migraineurs were found to experience aura, based on a diagnosis made from their own reports (Schürks, Buringm, and Kurth, 2010), and around 32% of MA sufferers report visual aura symptoms (Steiner et al., 2003). As vision is of interest here, this will be the focus, although it is important to remember that the aura is highly variable, and not exclusive to vision.

There is a huge variety in the visual aura symptoms reported, however, there are some recurring trends despite the variations. Many patients report their aura to typically begin as a small geometric pattern, known as a fortification pattern, somewhere in the visual field. This expands over a short period of time, around 20 minutes is a typical time period, to cover large areas of the
visual field. The fortification patterns surround a central scotoma, or hole in the visual field, which appear to scintillate: there is a certain shimmering effect associated with them. Illustrations of these patterns can be seen in Sacks (1992); Schott (2007) and Lashley (1994). Whilst this is a common pattern of aura, it must be stressed that there is a wide variety of disturbances of vision that have been reported by the patients of Sacks (1992). Other distortions reported by these patients include geometric patterns such as tunnels and spirals. Phosphenes, which are apparent flashes of light, and additional visual disturbances including micropsia and macropsia, the feeling that everything is much smaller or much larger respectively, are also commonly reported by these patients. Other sufferers studied by Sacks (1992) have reported distortions of vision that become progressively more disjoint, similar to a Cubist painting.

The rate of advancement of the aura across the visual field is of particular interest. The speed of these visual field disturbances matches the speed of advancement of a phenomenon called cortical spreading depolarisation (CSD) (Lauritzen, 1994), leading researchers to believe that there may be an association.

Cortical Spreading Depolarisation

Early work conducted by ao (1944), whilst studying epilepsy, found there to be a curious phenomenon that he termed 'cortical spreading depression' when the cortex was perturbed, either by mechanical, electrical or chemical means. When the cortex was disturbed, there originated a spreading wavefront of depolarisation - strong activity from the neurons at the wavefront. This strong activation of the wavefront itself was then followed by a period of inactivity, which ao (1944) termed 'depression', this state occurred after a latency of 10-40 seconds after the wavefront passed. This depression was unresponsive to any stimulation, but recovered spontaneously after a short period of time (around 5 to 10 minutes). The movement of the wavefront propagated across the cortex, but was confined to the hemisphere it originated in. It was possible to elicit CSD in many areas of the cortex, with the exception of Brodmann Area 12, part of the frontal lobe. It is unclear why this area was not susceptible to CSD.

The association with migraine is through the speed of propagation of this travelling wavefront. The migraine aura progressively engulfs more of the visual field as time goes on. A migraine sufferer himself, Lashley (1994) sketched the progression of the aura across the visual field, the expansion of the scototma, in a time course of around 7 to 10 minutes. The rate of expansion of the
aura was seen to correspond to that of the progression of the CSD across the cortex, around 2-6mm/min (Lauritzen, 2001). Thus CSD was associated with the migraine aura: the aura is thought to be caused by this wavefront of strong neural activity, and the scotoma is the depressed state of the cortex after the wavefront has passed.

Duckrow (1991) also reported an increase, followed by a decrease in blood flow in one hemisphere of the frontal cortex with CSD, as measured using the rat as an animal model. The rat was anaesthetised and blood flow measured using chemical tracer techniques and subsequent decapitation. CSD and increased blood flow are both associated with migraine; increase in BOLD response was found to increase with headache intensity, and also to progress at the same rate as CSD (Aurora, Cao, Bowyer, and Welch, 1999).

There have been suggestions that the pain of migraine is caused by vasodilation to accommodate this extra demand for nutrients to the brain to provide the materials for the excessive activity of the CSD wavefront (Bolay, Reuter, Dunn, Huang, Boas, and Moskowitz, 2002). Bolay et al. (2002) demonstrated that the vasodilation stimulates the trigeminal nerve. The trigeminal nerve is responsible for pain sensation in the face and head, including the cerebral blood vessels. Therefore, the throbbing headache of migraine could have its origins here. Although some nerves provide a fast, reflexive mechanism in response to pain, the trigeminal nerve is also involved in more complex noxious responses. The trigeminal nerve projects to the thalamus in order to relay noxious (pain) information to the higher cognitive areas than simply its own reflex mechanisms. Abnormalities in the responsiveness in trigeminal neurons in migraineurs result in a lower noxious threshold on heat stimulation, when compared to controls (Moulton, Burstein, Tulley, Hargreaves, Becerra, and Borsook, 2008). The trigeminal nerve also responds to light (Okatamoto, Tashiro, Chang, and Bereiter, 2010), suggesting a possible mechanism for the photosensitivity of migraine. However, it is unclear in which direction causality lies: vasodilation can also occur as a result of stimulating the trigeminal nerve (May, Buechel, Turner, and Goadsby, 2001). Therefore although this nerve is associated with the pain of migraine, it is unclear whether the dilation causes the nerve activation (and pain sensation) or whether the nerve activation causes the dilation.

The beta-blocker propranolol, which reduces high blood pressure, has been shown to help prevent migraine in some sufferers, especially when accompanied by behavioural measures; 70% participants showed clinical improvement
(a 50% reduction in migraine) in both MA and MO sufferers (Holyrod, Cot- 
trell, O’Donnel, Cordingley, Drew, Carlson, and Himawan, 2010). Propranolol 
acts to reduce cerebral blood flow, and cerebral perfusion pressure (Aqyagi, 
Deshmukh, Meyer, Kawamura, and Tagashira, 1976), therefore this could ac-
count for its effectiveness in treating the symptoms of migraine. There is also 
evidence to directly challenge the blood pressure hypothesis: Seçil, Uende, 
Beckmann, Oezerkan, and Başoğlu (2010) found no difference in migraineur 
blood pressure before, during, or after an attack, which would be expected 
if blood pressure were the issue. Obese people tend to have higher blood 
pressure (e.g. Tesfaye, Nawi, Minh, Byass, Berhane, Bonite, and Wall, 2007), 
therefore it might be expected that there will be more obese migraine sufferers 
than controls. However, Téllez-Zenteno, Pahwa, Hernandez-Ronquillo, Gacia- 
Ramos, and Velázquez (2010) found no association between BMI distribution 
in migraine sufferers compared to controls.

This suggests that the effectiveness of betablockers might lie in another of its 
effects. As propranolol is a small molecule it is able to cross the blood-brain 
barrier itself and therefore affect many other systems, such as the noradrenaline 
system, and serotonin systems. Shields and Goadsby (2005) showed that pro-
pranolol normalised thalamic activity induced by stimulation of the superior 
sagittal sinus. However, the authors attribute this to propranolol inhibiting 
the thalamocortical response directly.

The repercussions of excess blood flow might mean that migraine is not as 
benign as previously thought. Schürks et al. (2010) found in their longitudi-
nal survey that women with MA were twice as likely to suffer stroke as MO 
sufferers. Other researchers have suggested links between migraine and lesions 
in the cortex (Tietjen, 2004). Additionally, Basser (1969) noted the similarity 
of the association with CSD in both epilepsy and migraine. Epilepsy is a 
neurological disorder, characterised by excessive cortical activity that spreads 
from a focal point of the seizure, typically a lesion. A similar mechanism 
might therefore be the cause of the spread of uncontrolled neural activity in 
both disorders. Thus there is some uncertainty as to whether the increased 
blood flow is primary or secondary to the CSD: it has been argued that CSD 
causes increased blood flow by demanding more metabolic resources (Coppola, 
Pierelli, and Schoenen, 2007). It has also been argued that the CSD is a result 
of the plasma leakage out of the blood vessels, caused by the increased blood 
flow (Buzzi and Moskowitz, 2005). However, increasing evidence suggests that 
in migraine, CSD is primary, and the blood flow is a response to it (e.g. Bolay 
et al., 2002). Of particular interest here is the fact that CSD is readily elicited
10.3. MIGRAINE AURA

in Brodmann Area 17, or the primary visual cortex (Lauritzen, 2001).

Sensitivity to Light

Light is a common trigger factor, affecting 38.1% migraineurs (Kelman, 2007). Vanagaite, Pareja, ren, White, Sanc, and Stovner (1997) showed that on increasing light intensity 87% of migraineurs reported pain induced by this stimulus, with lower pain thresholds than controls. There was also an association between typical attack duration and pain threshold; the shorter the attack tended to be, the more light the individual could tolerate. Okatamoto et al. (2010) suggested that the pain could originate in the trigeminal ganglion, which responds to light; this might account for some of the photosensitivity in migraine.

The degree to which migraineurs are disturbed by visual stimuli depends on other properties of the stimulation, such as wavelength. Main et al. (1997) found not only that migraineurs had generally lower pain thresholds than both tension headache sufferers and controls, but also that medium wavelength (398-503nm, reddish light) light proved to be the most effective at causing discomfort for migraine sufferers, whilst controls had higher tolerance thresholds for these wavelengths. In addition to this, fMRI evidence shows that removing certain wavelengths can have a beneficial effect: elevated cortical activity was normalised by viewing patterns through a correctly coloured filter specific to that particular individual (Huang et al., 2011). It is important to note that this was not merely due to a reduction in light intensity caused by using the filter, as the correct coloured filter reduced the discomfort thresholds and activation more than a grey filter. The coloured filter had to be optimal for that particular individual to see the benefits; a filter of a different colour did not have the same effect in reducing elevated cortical activity. The elevated activity seen in migraineurs compared to controls was initially elicited on viewing striped patterns, indicating that the spatial properties of the stimulus are critical.

As well as luminance and wavelength, contrast also affects how disturbing migraineurs find the stimulation. Mulleners, Chronicle, Palmer, Koehler, and Vredeveld (2001) reported that those suffering from MA and MO both had lower thresholds for pain induced from a visual stress test stimulus. The test stimulus consisted of a square-wave grating, of varying contrasts. Higher contrasts were tolerated by controls than by migraineurs.

Additionally, the spatial and temporal attributes of the stimulus also affect
the degree to which migraine sufferers find light stimulation uncomfortable. Striped patterns, as well as flickering lights, are well-known for being difficult stimuli for migraineurs, and can induce migraine attacks (Shepherd, 2000). Marcus and Soso (1989) reported that stripes caused discomfort in 82% of migraineurs, compared to 18% of controls. This suggests that migraine sufferers are particularly susceptible to certain patterns of visual stimulation. The spatial frequency of the pattern is crucial to its ability to cause discomfort in migraine sufferers; striped patterns of around 3 cycles/degree have been shown to be particularly problematic, compared to both lower (0.5 cycles/degree) and higher (12 cycles/degree) frequencies (Harle and Evans, 2006). This increased sensitivity is accompanied by increased cortical activity compared to controls; Huang, DeLano, and Cao (2006) presented fMRI evidence to suggest that migraineurs show elevated cortical activity in response to 1.2 cycles/degree gratings. The experiment by Huang et al. (2006) was conducted at a lowered luminance (around 10 cd/m$^2$) than previous studies reporting a higher spatial frequency for the peak in discomfort judgements (e.g. Wilkins et al., 1984). Therefore it might be argued that the shift in peak discomfort judgements could be due to the shift in the peak in the contrast sensitivity function to lower spatial frequencies with lower luminance. However, the peak of the contrast sensitivity function does not shift down to 1.2 cpd until much lower luminances, of around 1.7 cd/m$^2$. The peak of the contrast sensitivity function is still approximately 3 cycles/degree at 10 cd/m$^2$, therefore lowered luminance is not a plausible explanation for the shift.

This suggests that migraineurs show particular sensitivity to certain patterns of light. It has been suggested that this might have a cortical origin: migraineurs show elevated cortical responses compared to healthy individuals, which results in both the excessive sensitivity to light stimulation and might be responsible for the migraine attack.

### 10.4 Hyperexcitibility

It has been suggested that one of the causes of migraine might be excessive cortical activation on sensory stimulation (e.g. Aurora et al., 1999). Excessive cortical activity could arise from various sources. One that has been suggested is cortical ‘hyperexcitibility’; migraineurs’ neurons responding more readily to stimulation than those of control subjects (Aurora et al., 1999). There is some confusion in the literature about what is meant precisely by hyperexcitibility. Coppola et al. (2007) use the strict sense of the term ‘hyperexcitibility’ to mean...
a response that is larger than normal to a stimulus, or a response to a stimulus that cannot normally be detected. However, other researchers (e.g. Aurora and Wilkinson, 2007) have a broader definition of hyperexcitibility, meaning an increased response, for a variety of potential explanations, including noise. Hyperexcitibility will be used here in the stricter sense, to mean a higher responsiveness of neurons to signals, with no change in noise levels compared to control subjects. This can be clarified using Equation 10.1:

\[ O = g(S + N_s) + N_i \]  

(10.1)

Encoding model equation from Zhaoping (2006). \( O \) is the output (response), \( S \) is the response to the signal, \( N_s \) is the signal noise, that are other neurons responding also due to the signal presentation, but do not contribute to the response (e.g. those responding to a mask), \( g \) is the gain control to amplify the response to the signal, and \( N_i \) is the internal noise, the spontaneous background response, irrespective of stimulus. In this case, greater \( g \) with constant \( N_i \) would represent hyperexcitation.

One possible mechanism for this excessive activity could be imbalance in the ions responsible for the creation of action potentials. An action potential (or spike) is caused by rapid ion transfer over the cell membrane. With a small amount of depolarisation, potassium ions leave the cell, restoring the membrane potential to its resting potential of around -70mV. However, on exceeding a certain threshold activation level, the voltage-dependent sodium gates open, and sodium ions are allowed into the cell. The massive influx of sodium anions and efflux of potassium cations changes the potential of the membrane massively. This propagates along the length of the cell as this change induces the next section of the cell membrane to open the voltage-dependent sodium channels. It has been suggested that the CSD originates in imbalances in potassium levels (Basser, 1969; Lauritzen, 1994). Evidence for this suggestion from Bogdanov, Multon, Chauvel, Bogdanova, Prodanov, Markarchuk, and Schoenen (2011) showed that introducing a high concentration of potassium ions induces CSD in rats.

Bogdanov et al. (2011) also showed that the drug lamotrigine suppressed
10.4. HYPEREXCITIBILITY

Figure 10.1: A diagram of a cell

The cell membrane showing ion transfer, and the synapse.

this induced CSD. Lamotrigine blocks sodium channels, and in so doing inhibits the release of the neurotransmitter glutamate. The action of glutamate is widespread throughout the brain, and its effects are generally excitatory. Glutamate inhibition might therefore provide relief from excessive activity. This is why lamotrigine is prescribed as an anticonvulsant in disorders such as epilepsy; the convulsions are thought to result from excessive activity. Interestingly, when used as a treatment for migraine, lamotrigine was found to have a beneficial effect only for MA, but not for MO sufferers (Manzoni and Torelli, 2008). This suggests not only that there might be a qualitative distinction between these two subsets of migraine, but also that there is a link between excessive activity, CSD, and migraine.

**Electrophysiological responses**

Early investigations into electrophysiology by Slatter (1968) showed EEG abnormalities in migraineurs on 20Hz flickering light stimulation using an oscilloscope. While there were large individual differences in these data, many of the migraineurs showed irregular rhythms in response to photic stimulation over the occipital areas, in particular the slow wave response (2-5Hz).

Electroencephalographic (EEG) techniques have been used to investigate cortical activity in migraine sufferers. The EEG response measures the change in electrical potential of groups of neurons, reflecting their activity in response to a stimulus, for example. Typically, this begins with a small positive response,
10.4. HYPEREXCITIBILITY

then a negative component, and then a large positive response. EEG data are divided into components denoting polarity and time of the response, thus N200 is a negative component at 200ms.

If migraine is the result of excessive cortical activation in response to a stimulus then this would be expected to manifest itself in electrophysiological data. This has been shown to be the case: on stimulation with light, there are higher amplitude EEG responses to visual stimuli (visual evoked potentials, VEP) in migraineurs compared to controls (Schoenen, Wang, Albert, and Delwaide, 1995; Lahat, Nadhir, Barr, Eshel, Aladjem, and Bistritze, 1997). Oelkers, Grosser, Lang, Geisslinger, Kobal, Brune, and Lotsch (1999) showed that elevated cortical responses to a low spatial frequency chequerboard pattern were accompanied by more illusions and discomfort reported by the migraineurs. Additionally, there was a longer N200 latency in migraineurs compared to controls, leading the authors to suggest that the attack might be due to an elevated, prolonged cortical response.

Other researchers report more extreme neural responses in migraineurs compared to controls for red/green chequerboard stimuli (Aurora et al., 1999), reversing patterns (Bohotin, Fumel, Vandenheede, Gérard, Bohotin, de Noordehout, and Schoenen, 2002), flickering square-wave gratings (Shibata, Yamane, Iwata, and Ohkawa, 2005), and also in response to auditory stimuli (Ambrosini, Rossi, de Pasqua, and Schoenen, 2003). Using MEG it was found that migraineurs show elevated N200 amplitude, which correlated with the frequency of migraine attacks: the greater the N200 amplitude the more often migraine attacks occurred in that particular individual (Lang, Kaltenhäuser, Neudörfer, and Siedler, 2004). Additionally, fMRI data also suggest that responses to gratings are greater in migraine sufferers compared to controls, especially for low spatial frequency stimuli of around 1.2 cycles/degree (Huang, Cooper, Satana, Kaufman, and Cao, 2003).

Other methods have also found evidence in support of hyperexcitation in migraineurs. Transcranial magnetic stimulation (TMS) is a relatively new technique of studying brain function. This technique uses a strong magnetic field to temporarily disrupt electrical activity in the brain. It can be used to both stimulate and inhibit function temporarily, depending on the parameters (e.g. frequency) of the stimulation (Sandor, 2007). TMS has been used in migraine research to investigate phosphene production. Phosphenes are perceived flashes of light in the absence of light stimulation to the eyes. Phosphenes are occasionally reported in migraine aura (Sacks, 1992) and can
10.4. HYPEREXCITIBILITY

also be elicited with TMS of the occipital areas. While the same proportion of control participants and migraineurs experience phosphenes in stimulation with TMS (Mulleners et al., 2001), migraineurs have lower phosphene thresholds (Aurora et al., 1999). This suggests that the cortex of migraineurs is more readily stimulated to this level than controls, perhaps as it is hyperexcitable.

Mulleners et al. (2001) highlight that it is sometimes difficult to induce phosphene perception: the TMS must be in the optimal location for a particular individual in order to be successful. Due to the variation in anatomy, this is difficult to ensure, thus the authors caution the interpretation of results using this technique - it might be the case that those not reporting phosphenes were simply not receiving stimulation in the optimum location. This is important as it is presented by the authors as a possible reason for the discrepancies between results of different studies. It is unclear why this would cause a difference between clinical and non-clinical test groups, however: it might require that the experiment be conducted so that the researcher is unaware of the group each observer belongs to, to guard against inadvertent sources of bias.

Evidence Against Hyperexcitibility

The electrophysiological data presented make a case for elevated cortical responses in migraineurs compared to controls. However, this is not unanimous; Khalil, Nicotra, and Wilkins (2010) showed there to be a reduced P100 response on the side of the head where the headache usually was to be found. In the accompanying psychophysical task (4AFC (four alternative forced choice) task to detect a vertical grating), reduced contrast sensitivity was found with the same lateralisation in migraineurs compared to controls. Other behavioural evidence does not support the hyperexcitation hypothesis. Elevated responsiveness to stimuli, with constant levels of noise, in migraineurs would be expected to result in heightened sensitivity, and therefore lowered thresholds for visual tasks. However, evidence suggests that migraineurs show higher thresholds for contrast sensitivity for gratings (Shepherd, 2000), particularly for low spatial frequencies (Benedek, Tajti, Janáky, Véecsei, and Benedek, 2002) and for flickering stimuli (McKendrick, Vingrys, Badcock, and Heywood, 2000). Therefore not all evidence supports the case that there is simply a stronger response to stimulation in migraineurs. One possible alternative that has been proposed is that the excessive responses are due to faulty inhibitory processes in migraineurs.
10.5. A LACK OF INHIBITION

Centre-surround architecture of ‘ON’ centre cells for the LGN cells. Right: ‘OFF’ centre cell, left: ‘ON’ centre cell. A stimulus in front of the ‘ON’ centre cell shown in grey. There will be a positive response from the ‘ON’ centre, but an inhibitory response from the parts of the stimulus that fall in the ‘OFF’ surround.

10.5 A Lack of Inhibition

Inhibitory processes could account for some of the deficits found in migraineurs compared to controls. Consider for example the ‘ON’ centre cells in the visual system (see Figure 10.2). Although there are both ‘ON’ centre and ‘OFF’ centre cells, the following logic applies to both, but with reversed polarity. The excitatory centre of the receptive fields of cells in both the LGN and the later visual cortex are surrounded by an inhibitory region. This inhibitory surround region is important for increasing the sensitivity to detect visual features. The particular feature a cell is optimised to detect, for example bars or edges, relates to the shape, orientation and polarity of the cell. In any case, uniform stimulation covering the whole of the cell will not elicit much of a response, as the input to the excitatory and inhibitory regions will cancel out (Hubel and Wiesel, 1962). Therefore, if there are deficits with the inhibitory surround, contrast sensitivity will be lowered.

Surround effects are only one possible form of inhibition, there are others, such as long range inhibitory connections. These are responsible for the sharpening of spatial frequency tuning - cells are more specific in their response than would be the case based on cell architecture alone. This extra selectivity originates in inhibition from other cells. This is particularly the case for oriented stimuli.
Deficits in inhibitory processing could account for a number of differences in visual processing in migraine. For example, Coleston, Chronicle, Ruddock, and Kennard (1994), asked participants to detect a striped target against a striped grating background of varying spatial and temporal frequency. Migraineurs showed an overall loss of sensitivity in comparison to controls. Additionally, peak sensitivity for spatial frequency modulations was shifted for migraineurs (both MA and MO) to around 5 cycles/degree in comparison with controls (around 2.5 cycles/degree). Peak sensitivity for variations in temporal frequency was around 5-8 Hz for the migraineur group, compared to 10 to 15Hz for the control group. The authors attribute their results to differences in connectivity in the precortical pathways. However, it could also be construed that this shift in contrast sensitivity could be accounted for by inhibitory differences in migraineurs affecting the shape of the contrast sensitivity tuning functions, as tuning relies on inhibitory processes.

The overall lowered contrast sensitivity would be compatible with explanations involving inhibition deficits. However, contrast detection thresholds are not specific enough to dissociate whether faulty inhibition is the problem in migraine, as this could also be accounted for by other explanations, such as excessive noise, reduced initial levels of excitation, or faulty neural synchronisation, which will be discussed in more detail below. Other psychophysical tasks, such as aftereffects and masking, are much better suited to assessing inhibitory function. Palmer, Chronicle, Rohan, and Mulleners (2000) used metacontrast masking to investigate inhibitory processes in migraineurs. Metacontrast masking is when the apparent contrast of a target is reduced by flashing flanking stimuli shortly after the target (Alpern, 1953). This effect can impede the detection of the target, and occurs by a masks presented shortly before or after the target (Macknik and Livingstone, 1998). Varying stimulus onset time affects the size of the effect: 100ms after target offset is the most effective time for backward masking to occur (Alpern, 1953; Macknik and Livingstone, 1998). It is thought that backward masking occurs because the transient signal of the target is inhibited by the sustained signal from the mask. The specificity of the timing effects indicate the disruption of the signal relies on the inhibitory processes, which take some time to work. Palmer et al. (2000) showed that MA sufferers were less susceptible to the metacontrast masking than controls, indicating that their inhibitory processes are lacking. Sodium valporate, which
is known to affect the GABAergic system, and therefore thought to affect inhibitory processes, was shown to normalise the results.

Reduced inhibition would also be expected to affect CSD. Dahlem and Hadjikhani (2009) used chicken retina to explore some of the properties of CSD and to test a mathematical model of this phenomenon. Their model includes two parameters that affect how easily the CSD initiates and propagates across the cortex. One parameter relates to the level of inhibition created by the neurons, and the other how much influence they have over one another, termed coupling strength. By adjusting these parameters, it could be shown that the model was a good replica of the behaviour of the CSD wave evoked in the chicken retina, in terms of susceptibility to CSD, and how the wave propagates. However, what these two parameters might relate to is still unclear; these parameters could be composed of multiple components that influence coupling and inhibition. However, it seems clear that there is a necessity for weak coupling, so there must be a level of interaction between the cells, but not that the action of one cell necessarily evoked action in others. Also there is a necessity for a certain level of inhibition to initiate the CSD. Therefore, as the CSD was affected by the inhibitory strength parameter, one suggestion is that migraine might be the result of faulty inhibitory processes, according to the Dahlem and Hadjikhani (2009).

**Mechanism of Inhibition**

One of the possible mechanisms of inhibitory processing revolves around the neurotransmitter gamma-amino-butyric acid (GABA). This neurotransmitter tends to have inhibitory effects, and this might be the root of the problem in migraineurs. For example, Carrera, Stenirri, Ferrari, and Battistini (2001) used a mouse model to investigate abnormalities with the calcium channel: a defect on a particular gene resulted in an abnormal calcium channel expression. The calcium channel is important for the release of the neurotransmitter GABA, therefore the results provide a potential mechanism to account from migraine: that the inhibitory system might be flawed.

Advances in drug therapy can also give insights into the causes of the CSD. Bowyer, Mason, Moran, Tepley, and Mitsias (2005) provided evidence that sodium valporate may make individuals less susceptible to CSD. This is a drug commonly taken to prevent migraines, which acts by increasing the level of the inhibitory neurotransmitter GABA in the brain.
Role of Serotonin in Overresponsiveness

Another possible mechanism for a lack of inhibition is neurotransmitter imbalances. Unlike ion channel mechanisms, neurotransmitters bridge the gap between neurons, allowing one neuron to induce activity in the next neuron, see Figure 10.1. The neurotransmitter serotonin has been suggested to be one possible candidate: insufficient serotonin might result in excessive activity. Serotonin is thought to act on GABAergic interneurons to inhibit glutamergic activity in some areas of the brain (Stutzmann and LeDoux, 1999). A lack of serotonin would result in less inhibition, and therefore more overall activity.

The level of serotonin is affected by exposure to light: excess serotonin has been found after sleep deprivation in bright light conditions in hamsters (Grossman, Mistlberger, Antle, Ehlen, and Glass, 2000). Serotonin is increased in bright lighting conditions, and bright light is needed to synthesise serotonin, thought to be one of the mechanisms of Seasonal Affective Disorder (SAD) (Danilenko, Putilov, Russikikh, Duffy, and Ebbesson, 1994). Rare photosensitive skin conditions have been shown to be affected by drugs affecting the level of serotonin (Doffoel-Hantz, Boulitrop-Morvan, Sparsa, Bonnetblanc, Dalac, and Bédane, 2009), strengthening the association between light and serotonin.

It is possible that insufficient levels of serotonin are connected to migraine. There has been shown to be a beneficial effect of SSRIs (selective serotonin reuptake inhibitors, serotonin agonists): excessive responses in migraineurs could be reduced to normal limits with SSRIs (Lang et al., 2004). Tryptophans are amino acids that increase serotonin levels, and evidence suggests that tryptophan depletion has a role in migraine: Drummond (2006) gave participants an amino acid drink, which either contained a balanced amount of all amino acids, or was deficient in tryptophans. By increasing light levels and measuring the glare pain ratings, the authors found that patients with depleted tryptophan levels were more susceptible to pain from the bright light than those whose were at normal levels.

Evidence for the Inhibition Hypothesis

Orientation discrimination is a visual ability that might be detrimentally affected by faulty inhibitory processes. Tibber and Shepherd (2006) showed higher orientation discrimination thresholds to detect both the orientation of
10.5. A LACK OF INHIBITION

stimuli and virtual stimuli in migraineurs (both MA and MO) compared to controls. Virtual stimuli were two small circles opposite each other. It would be possible to imagine a line connecting them, and the orientation of this imaginary line was judged. There was no difference in reaction time or contrast detection levels. However, the elevated thresholds were for oblique stimuli only, cardinal orientations showed no such deficit. This could be accounted for by a lack of inhibition, which is known to sharpen orientation discrimination (Seriès et al., 2004). However, why this detriment should only be for oblique, not cardinal, orientations is unclear.

Another way to examine cortical inhibitory processes is through a phenomenon called the cortical silent period. This is a period of inactivity following TMS of the motor cortex, thought to be controlled by inhibitory neurotransmitters GABA and dopamine, which have associations with inhibitory processes. It has been demonstrated that the cortical silent period is shortened in migraineurs compared to controls, suggesting that there is less of this inhibition in this group (Aurora et al., 1999).

However, not all studies are unanimous: other studies find that motor cortex inhibition was normal in migraineurs (Siniatchkin et al, 2007). The discrepancy in results could be due to technical details, or differences in migraineur type: the Aurora et al. (1999) group was exclusively composed of MA, whilst Siniatchkin, Averkina, Andrasik, Stephani, and Gerber (2006) tested exclusively MO sufferers. Additionally, Lang et al. (2004) used MEG techniques to investigate N20 and P35 somatosensory responses. N20 is associated with excitation, P35 with inhibition of this area. A mixture of 15 MO and 15 MA showed higher N20 amplitude than controls, indicating increased excitatory activity, however P35 was no different from controls, indicating normal levels of inhibition. More specificity with respect to migraineur subtype might have provide greater insights: it could be that inhibition differences in MA might have been diluted by the inclusion of MO participants. However, as it stands, results indicate that inhibition failure alone cannot entirely account for differences between all migraineurs and controls.

Additionally, McColl and Wilkinson (2000) specifically investigated inhibitory connections in migraineurs compared to controls using delayed onset masking effects. A Gaussian target was hidden by a sine grating mask, presented either simultaneously (baseline) or with a delayed onset time. As there is a slight time delay for the inhibitory connections to work, it would be possible to see if there was reduced inhibition, as this would be differentially affected by the
onset time of the mask. Although greater overall masking effects were found for migraineurs than controls, the delay time made no group difference to the amount of masking. This suggests that inhibitory connections are normal in migraine populations. Using similar logic, Huang et al. (2006) investigated the detection of a grey bar with a mask of progressively delayed onset time. The authors found no difference in error rate for target detection between MA and healthy controls for the different onset times. This indicates that inhibitory processing is intact in the clinical group.

Psychophysical studies involving various after-effects have provided evidence that migraineurs have intact inhibitory processing. After-effects present themselves after a time of viewing a certain pattern, after the stimulus is removed. For example, after a time of viewing a tilted line, the stimulus is replaced with a vertical line. This vertical line is perceived by the observer to be tilted in the opposite direction (Gibson and Radner, 1937). This is suggested to be a consequence of inhibitory interactions; during adaptation, the cells signalling vertical gratings inhibit those signalling other orientations around them. After the adapting stimulus is removed, there is a removal of this inhibition, ‘disinhibition’, which results in a relative increase in firing rate from these previously inhibited orientations, and thus they are perceived (Georgeson, 1976). Increased disinhibition will result in increased after-images. However, if inhibition is faulty, there will be less ‘disinhibition’ and a weaker after-effect. After effects occur in many visual domains as well as tilt, including motion: after adapting to motion, stationary stimuli appear to drift in the opposite direction to the adapting motion. This particular example is known as the ‘waterfall illusion’, although there are many other types of after-effect. (Shepherd, 2001) showed that migraineurs have longer after-effects, both of motion and tilt, than controls. (Shepherd, Palmer, and Davis, 2002) also showed that the increased duration of the tilt illusions is not accompanied by a change in sensitivity to orientation differences, indicating that this is not due to sensitivity differences between migraine and control groups (Shepherd et al., 2002).

Battista, Badcock, and McKendrick (2010) failed to find differences in the strength of the motion after-effect for a Gabor stimulus moving against a grating. In this experiment the size of the effect was measured by the observer reversing the direction of the grating motion until the apparent motion of the after-effect was cancelled. Additionally, Battista, Badcock, and McKendrick (2011) conducted a subsequent experiment based on the Chubb illusion. In the Chubb illusion a stimulus against a dark background appears lighter than an identical stimulus against a white background. This is thought to be explicable
through suppression effects; the cells signalling a lighter background suppress the output of the ones signalling the target stimulus, and vice-versa. As the Chubb illusion is based on surround suppression effects, and the greater the surround suppression (from the lighter background), the greater the reduction in perceived contrast of the target stimulus. Faulty inhibitory processes would be expected to manifest as a less powerful Chubb illusion in the clinical group. Battista et al. (2011) showed that there was actually a greater effect of the Chubb illusion for migraineurs than controls, the opposite of the prediction. Therefore from these two studies it appears that the inhibitory facility of migraineurs is intact, and therefore this hypothesis cannot account for all migraineur differences of visual function.

There are some indicators that there might be inhibitory deficits in migraineurs that can be inferred from pharmacological treatments, such as imbalances with the neurotransmitter GABA, which is associated with inhibitory processes. However, drugs have far reaching effects, therefore caution must be exercised when using pharmacological evidence to infer deficits in migraineurs, as an association is inconclusive. Some behavioural and electrophysiological evidence supports the possibility of faulty inhibition, however, the results are not always unanimous. What is evident from studies of visual function is that there are specific differences in migraineurs compared to controls. There is evidence of lowered contrast sensitivity, greater masking effects, and stronger motion, tilt and colour after-effects in migraineurs compared to controls. Although these could be associated with faulty inhibition, there are some aspects, such as inconclusive evidence and lack of temporal variation, which cannot be accounted for by faulty inhibition. Therefore, another explanation has been proposed, this one involving the level of noise in the cortex.

10.6 Elevated Internal Noise

Many of the electrophysiological and behavioural findings mentioned previously could be accounted for in terms of excess cortical noise in migraineurs compared to controls. A higher general level of cortical activity could be the result of excessive noise - more active neurons would appear as higher levels of electrical activation, as well as showing a higher metabolic expenditure, thus being manifest in EEG, fMRI and PET scans. This would be a greater $N_i$ from the encoding model (see Equation 10.1).

Excessive cortical noise in migraineurs can not only account for increased
baseline levels of activity, but also explain the increased amplitude of EEG responses to stimuli; increased noise would mean that migraineurs might show an increased difference in peak and trough of the EEG components as they respond more strongly in order to boost the signal through gain control. Visual performance differences could also be explained by excessive noise. Higher contrast detection thresholds would be seen if the observer could not discriminate the signal from their own internal background noise as readily as controls. Greater masking effects might be seen with more cortical noise - again, the level of cortical noise might be combined with the external noise of the mask to make the signal harder to detect for the clinical group compared to controls.

Ditchfield, McKendrick, and Badcock (2006) investigated the effect of internal noise in migraineurs using global form and global motion tasks. Global form tasks involve detecting a shape, for example a spiral, composed of coherently placed dots. Noise is added to the signal by introducing other dots in random locations. In this experiment, observers indicated the direction of the spiral. Similarly, global motion tasks involve coherent signal dots moving in the same direction, whilst noise dots move randomly. Both these tasks rely on good integration of signal information over space or time, whilst ignoring noise. Greater internal noise would make the integration task more difficult, and therefore a greater percentage of signal dots compared to noise dots would be needed in order to perform the task. These authors found both global motion and global form to be impaired in migraineurs compared to controls, suggesting that excess noise might be a problem for the clinical population.

However, McKendrick, Vingrys, Badcock, and Heywood (2001) found evidence of specific visual impairments that would not be predictable from overall elevated internal noise levels. McKendrick et al. (2001) applied a range of visual tasks to assess the visual performance of migraineurs compared to controls, including contrast sensitivity, orientation discrimination, motion detection and discrimination, and global motion detection. These tasks were designed to discriminate between the parvo and magnocellular pathways, in order to see which processing stream might be responsible for the deficit in migraine. The early visual system is thought to be divided into two main processing streams, the parvo- and the magnocellular pathways. The magnocellular pathway deals with low spatial frequency, high temporal frequency information, whereas the parvocellular system is responsible for high spatial, low temporal stimuli. By presenting a carefully chosen set of stimuli and tasks, the relevant pathways can be discriminated by exploiting their respective properties. For example, the global motion task was used here to target the magnocellular system. In
10.6. ELEVATED INTERNAL NOISE

In this global motion detection task, migraineurs were found to have higher motion coherence thresholds than controls, both for the centrally and peripherally presented tasks. In total, migraineurs performed more poorly than controls for three out of six tasks designed to assess magnocellular function: orientation discrimination at low spatial frequency (0.5 cycles/degree), and global motion discrimination both presented centrally and peripherally. This suggests that there might be a magnocellular deficit in migraineurs, however, given that the deficits were not seen on all magnocellular tasks, it is a weak argument. However, although weak, this evidence supports the argument that increased noise might be present only in one specific pathway, e.g. the magnocellular pathway. However, subsequent research showed that the evidence for a purely magnocellular deficit is not conclusive (McKendrick and Badcock, 2003).

Visual search tasks could be a good test of effects of internal noise in migraineurs: pop-out visual search tasks require the participant to detect a target against a background of distracting stimuli. If migraine sufferers were subject to excessive noise in the cortex, then this should predict poorer visual search performance, possibly from an inability to discriminate the target from the distracters. However, evidence suggests that this is not the case in migraine. Wray, Mijovic, and Kosslyn (1995) conducted a study using a pop-out visual search task, and demonstrated that migraineurs are no worse than controls. This finding has also been replicated: Conlon and Humphreys (2001) showed migraineurs to be no different from controls at a visual search task. By contrast, a group of migraine-free, high visual discomfort sufferers (classified using the Conlon et al. (1999) Visual Discomfort Scale) were slower at the task than the controls. This indicates that the test is not so insensitive that such differences cannot be detected.

In addition to this, longer after-effects and increased surround suppression are difficult phenomena for internal noise explanations alone to account for. According to Mather (1980), MAE is based on the integration of outputs of detectors that respond to all directions. On viewing a stimulus, the adapting neurons signal more strongly, and their output is depressed on removal of adapting stimulus, meaning that the sum output, and therefore overall motion percept, is in the opposite direction. Wagner, Manahilov, Loffler, Gordon, and Dutton (2010) showed that not only is there more noise in migraineurs compared to controls, but this noise is compounded by responses to the signal. These researchers assessed performance in migraineurs compared to controls using a target disk hidden in a Gaussian noise mask. This psychophysical performance was modelled using a variant of the perceptual template model.
Parameters critically involved include additive noise (a constant amount of noise added to the performance prediction) and multiplicative (noise scales with signal strength). Behavioural evidence was consistent with the multiplicative noise model: migraineurs were not only worse at the task than controls, but the difference between the groups increased with higher signal trials. This suggests that the greater the signal, the more the noise is generated, and the greater the detriment to performance. This might be more likely to be created by inhibitory problems than simply differences in baseline levels of neural activation.

Clearly it is difficult to account for the performance differences between migraineurs and controls. Many experimental outcomes fail to replicate and are not found universally across migraineurs. This could be due to a variety of factors, for example differences in exact stimulus used, or migraineur type. It has also been suggested that the apparent overresponsiveness is due not to an increased level of activation, but a decreased level of initial activation that does not habituate in the way typical of controls (e.g. Coppola et al., 2007). This will be discussed in the next section.

10.7 Habituation Deficits

One possible cause that has been suggested is that migraineurs do not show the same levels of habituation as controls. Habituation is the adaptation that takes place with repetitive stimulation: after repeated stimulation, the response to a particular stimulus is diminished. Thus a tolerance is developed: stronger stimulation of this kind is needed to elicit the same initial response magnitude. After stimulus cessation, this habituation level will gradually return to normal, so that if the stimulus appears after this time, the response will be as high as initially once more. It has been suggested that this process might be faulty in migraineurs.

Evidence for a lack of habituation in migraineurs has been shown in the EEG responses between migraineur and control groups, in particular in a component called Contingent Negative Variation (e.g. Ambrosini, Magis, and Schoenen, 2011). Contingent Negative Variation (CNV) is a slow wave EEG component. CNV is the electrophysiological response seen to a stimulus that follows a warning indicator. For example, a warning tone appears, and then the participant responds to a series of flashes of light (e.g. Mulder, Linssen, and Passchier, 2001). This is seen as the electrophysiological reaction to a novel, but ex-
HABITUATION DEFICITS

pected, stimulus. The early CNV component has been found to be increased in amplitude in migraineurs compared to controls, and this component does not habituate to a similar extent in migraineurs compared to controls (Kropp and Gerber, 1994).

Migraineurs do not show the same pattern of function and EEG responses all the time. Results tend to vary over an attack cycle, tending to be worst in the days shortly preceding the attack, and deficits minimised shortly after an attack (Sacks, 1992). CNV is also associated with the time course of the migraine cycle - CNV amplitude is greatest, and habituation least, in the days shortly preceding the attack itself (Siniatchkin et al., 2006). It is worthwhile noting that the CNV amplitude does not vary with the behavioural performance of the individuals; no differences were found between migraineur and healthy volunteer reaction times, although the migraineurs did report that it was more effortful to complete the reaction time task in the days preceding their attack (Mulder et al., 2001). Migraineurs showed elevated CNV response amplitude, and a reduced habituation of this response over subsequent trials compared to controls. This reduced habituation has been suggested to be the cause of the migraineur anomalies.

Early work investigating this possibility showed there to be no lack of habituation in migraineurs compared to controls. In fact, Afra, Cecchini, de Pasqua, Albert, and Schoenen (1998) found that migraineurs habituate to visual stimulation in the form of a chequerboard after around 12 minutes, whereas controls did not.

However, other researchers have found evidence for deficits in habituation in migraineurs. Bohotin et al. (2002) investigated VEPs induced by pattern reversal (PR-VEP) in migraineurs compared to controls. The PR-VEP responses of migraineurs were shown to have less habituation than controls. Participants were then subjected to TMS, and VEPs were measured again. After stimulating migraineurs with 10Hz excitatory TMS, their PR-VEP showed habituation similar to that of healthy volunteers. Migraineurs showed no change with 1Hz inhibitory stimulation. These authors suggested that this was potentially due to reduced baseline level of activation in migraineurs compared to controls, which was normalised by the 10Hz excitatory stimulation. The authors claim that this lowered preactivation level is the cause of the reduced habituation. The apparent paradox could be accounted for by the work of Schoenen (2006). The effects of TMS depend on stimulation frequency: TMS can be excitatory or inhibitory, depending on the frequency of stimulation. Schoenen (2006)
10.7. HABITUATION DEFICITS

showed that after repetitive TMS at a frequency that usually excites healthy volunteers, inhibition was found in migraineurs. This suggests that the rTMS normalises faulty inhibitory processing in migraine sufferers, but as it is the inhibition systems that are faulty, it makes no difference to migraineurs when inhibitory stimulation is given.

Electrophysiological abnormalities are not restricted to those with migraine diagnosis: Siniatchkin, Kropp, and Gerber (2001) investigated those with a family history of migraine. Those with a high risk of migraine, indicated by many close family members suffering from the disorder, showed similar EEG abnormalities as migraineurs. Notably, the early component (iCNV) showed increased amplitude and decreased habituation in migraineurs and the high risk group, but the low risk group did not show this. Therefore, it could be that the early CVN component could be used as an indicator of the disorder. Biological markers such as this are interesting as there might be a causal association, which is worth investigating further.

The lack of habituation in migraineurs compared to controls has been associated with imbalances in the neurotransmitter serotonin. Ambrosini et al. (2003) showed that whilst controls habituated to repetitive auditory stimulation, the same auditory potential increased in amplitude in migraineurs. Interestingly, it was found that the migraine abnormality could be reversed with the serotonin reuptake inhibitor fluoxetine. Additionally, Evers, Quibeldey, Grotemeyer, Suhr, and Husstedt (1999) showed a loss of P300 habituation in migraineurs in the interictal period, whilst habituation normalised during the attack. This study also showed decreased serotonin levels in the platelets of migraineurs compared to controls, and especially during the attack period. The level of serotonin in the platelets is not the instantaneous level in the brain, it takes time for the level of serotonin in the brain to get into the platelets. As there is a delay associated with measuring platelet levels of serotonin, this would correspond to levels being lowest just before the attack. This suggests that serotonin levels decrease until a threshold level, then normalise after the attack.

However, the results are not unanimous: Chen, Lin, Fuh, Hämäläinen, Ko, and Wang (2011) investigated PR-VEPs in migraineurs, and found increased P100 amplitude in MA sufferers. MO showed potentiation (increasing activation over subsequent stimulus presentations) with a low frequency, repetitive TMS stimulation. This does not suggest a lack of habituation resulting from a lower preactivation level of activity in these groups.
Lack of habituation can account for the prolonged aftereffects shown in migraine sufferers (Shepherd, 2001; Shepherd et al., 2002). Adaptation of neurons to a particular stimulus would mean that their firing rate would drop with prolonged viewing. Without this habituation, the adaption would be stronger and thus the decay period longer. However, lack of habituation cannot be the sole explanation for migraine as it does not account for the differences in contrast sensitivity thresholds, masking or surround suppression effects seen in migraineurs. The fact that some migraineurs show more efficient visual search performance can also not be explained by habituation - it would be predicted that a lack of habituation to novel stimuli might mean poorer search performance, as noise stimuli might continue to be salient to the individual. As the search performance of migraineurs has been shown to be better than that of controls (Wray et al., 1995) then there is little outside the electrophysiological evidence to bolster the habituation explanation.

10.8 Limitations of the Explanations

One of the problems with the theoretical explanations of migraine is that the distinctions between them are very fine. This relates to the difficulty in designing experiments that specifically target particular explanations and thus discriminate between them. Subsequently much of the evidence provided in support of any particular explanation does not exclusively test any one over any other explanation. It is also possible that these related explanations are part of the same underlying cause of migraine. A difference in the activity of any given neurotransmitter could provide a possible mechanism for more than one theoretical cause of migraine. For example, serotonin is a neurotransmitter with widespread effects. It is related to the 'fight or flight', sympathetic system activation. Therefore serotonin provides at least two explanations - overall hyperexcitation, and reduced habituation levels. Another example is GABA, which also has many effects. It has been suggested that a deficit in GABA could manifest itself in deficits in inhibitory processes, and also habituation deficits (Brighina, Palmero, and Fierro, 2009). Evidence tends to support both explanations, for example repetitive TMS up-regulates these GABAergic inhibitory circuits, restoring both inhibition and habituation at the same time. Thus carefully designed experiments are needed to target theoretical explanations, as well as ascertaining level of function in migraine.

Another big issue for the established explanations of migraine is that there
are discrepancies between the electrophysiological and behavioural data. For example, Mulder et al. (2001) showed that differences between migraineurs and controls are found in CNV amplitude, but not in reaction time on their behavioural task. One potential reason for this is compensation. For example, it might not be possible for migraineurs to adjust their neural responses, but they might be able to adjust strategy, or show learning effects in the behavioural response. This is difficult for any explanation of migraine to reconcile. Additionally, none of the above explanations can account for all of the electrophysiological and behavioural data. This could be explained in part by wide individual variation, and therefore it could be the case that some studies failed to make an important theoretical distinction between migraineur subtypes. It has been suggested that although MA and MO are similar, these are actually distinct disorders (Manzoni and Torelli, 2008).

One large problem with the psychophysical data is that the underlying issues with migraineur visual processing, such as global motion processing deficits, might in fact have more banal causes. Shepherd, Beaumont, and Hine (2012) used tests of global motion processing to investigate the possibility of increased noise in migraineurs compared to controls. Three types of noise were added to a random dot stimulus to determine relative motion discrimination thresholds. The first was static noise in the surround, the second static noise in the stimulus, and the third was dynamic noise. Dynamic noise was be thought to be a test for the possibility of stochastic resonance. Stochastic resonance suggests that there is an elevated baseline activity level (e.g. from the dynamic noise) means that the resting state of the neurons is closer to threshold. Thus only a small amount of stimulation is needed from the intended stimulus for perception. Effects of stochastic resonance were hypothesised to be greater in migraineurs, if their baseline activity level was elevated relative to controls. Such effects of dynamic noise were found, thresholds improved more with the addition of dynamic noise for the migraineurs compared to controls. However, when contrast sensitivity was included as a covariate in the analysis, all group differences disappeared. Therefore although there may be results suggestive of internal noise in migraine, this study highlights the importance of accounting for differences in contrast sensitivity of migraineurs compared to controls when assessing more complex visual processing.

A more complex explanation might be able to reconcile some of the differences in the data and provide a reason for some of the incongruencies, such as increased neural activity, but lowered sensitivity. One such explanation might lie in the synchronisation of neural responses in MA sufferers.
10.9 Synchronisation of Neural Responses

One key feature unaccounted for by the explanations described above is the origin of the CSD, which seems likely to have a role in propagation of the migraine attack. This could potentially be instigated by abnormalities in the rhythm of synchronised oscillations in the brain. There are two types of oscillatory activity commonly described as ‘synchronisation’. The first relates to the stimulus. Oscillations that are directly in response to the stimulus presentation (‘phase-locked’ to the stimulus), are termed event-related synchronisation. The second relates to synchronisation between oscillatory activity in different cortical areas. This is related to the stimulus, but not phase-locked to it - these oscillations follow a particular pattern, but the onset time can vary. It is these that are of interest here. Both types of oscillatory activity occur together when they appear, therefore the event-related synchronised activity will first be considered.

Event-Related Desynchronisation

Synchronised, event-related (phase-locked) neural activity is thought to be an important component of the ‘idling’ state of the normally functioning brain (Pfurtschneller and de Silva, 1999). The ‘idling’ state is when the brain is not stimulated or performing any particular task. Phase-locked oscillations are found in different frequency bands. The alpha band (9-24Hz) oscillations are thought to represent the idling state of the visual areas (extrastriate cortex), and are therefore of interest here.

In a recent review, Pfurtschneller and de Silva (1999) reports that synchronised alpha activity is thought to be the ‘idling’ state of the visual cortex. This can be seen by analysing the spectral content of the EEG activity. EEG activity can be analysed by averaging, so that the response that is phase-locked to the stimulus (the event-related response) can be disambiguated from the noise. However, by analysing the spectral content of the EEG response, it is possible to analyse oscillatory responses. The authors make the distinction that event-related synchronisation is identifiable as a peak in the amplitude spectrum of the EEG signal on the presentation of a stimulus, that was not there before. Event-related desynchronisation is a peak that was in the amplitude spectrum initially, that reduces to baseline in the presence of the stimulation. Critically, event-related synchronisation and event-related desynchronisation are phase-locked to the stimulus. Pfurtschneller and de Silva (1999) reports that the alpha band activity shows event-related desynchronisation on processing sen-
sory information. It has been demonstrated that this change in alpha band activity can drive higher frequency oscillations, such as gamma band oscillations (Pfurtschner and de Silva, 1999; Fries, 2009).

There is some evidence for desynchronisation during a migraine attack. The scintillations behave as a visual stimulus: evidence from Hall, Barnes, Hillebrand, Singh, and Holliday (2004) demonstrated desynchronised alpha and synchronised gamma activity in the rhythms of an EEG during the migraine aura. Phase-locked desynchronisation of alpha activity was mostly in the extrastriate visual areas and temporal lobe, and existed for the duration of the scintillations of the migraine aura (around five minutes duration). By contrast, desynchronised gamma rhythm was only found in the temporal lobe, and lasted longer than the desynchronised alpha activity, taking around 16 minutes to return to normal levels.

**Synchronisation between Cortical Areas**

Induced gamma band (25-80Hz) oscillations are of interest here. These oscillations are thought to play a role in synchronisation of activity in different cortical areas (Bertrand and Tallon-Baudry, 2000). These are not phase-locked to the stimulus directly, and therefore not the same as the event-related synchronisation and event-related desynchronisation previously discussed. Induced gamma activity is time-locked to the stimulus, but there is variation in precisely when this pattern of activity related to the stimulus actually begins. This could be because of how gamma synchronisation occurs. It is thought that gamma oscillations can synchronise activity across different regions of the cortex, and are thought to play a role in binding features into a coherent whole percept (Bertrand and Tallon-Baudry, 2000). For example, the same patterns of cortical activation have been found across the occipital and the temporal cortex, but this was dependent upon percept (Tallon-Baudry, Bertrand, Delpuech, and Pernier, 1997). In the study by Tallon-Baudry et al. (1997), coherent and incoherent patterns of dots were the stimuli: in one case it was possible to perceive a dog hidden in the dots, in the other condition there was no such percept. Alpha activity was the same whether or not the observer perceived a dog or not. However, induced gamma band synchronisation occured only when observers perceived the dog, despite the overall visual stimulation being the same.

Synchronisation depends on the strength of the coupling between cells, which is the ability of the cells to excite or inhibit other cells. Gray and McCormick
10.9. SYNCHRONISATION OF NEURAL RESPONSES

(1996) identified chattering cells, which are cells that fire in short bursts, and have the ability to induce other cells to synchronise to a particular rhythm. This synchronisation seems to be controlled by the balance between excitatory and inhibitory extracellular input (Wright, 2009). Excitation is needed to elicit initial activity from a population of cells: Gieselmann and Thiele (2008) showed that oscillations decrease in frequency with increasing size of stimulus, indicating that they need an excitatory input. The main drive of the gamma oscillations is when the stimulus is around 4.71 degrees of visual angle. This would project to 1mm in V1, which would be sufficient area to span more than one hypercolumn, and therefore impact units that do not have overlapping receptive fields. A repetitive stimulus with cycles of this size could be even worse, as multiple hypercolumns would be activated. The spatial frequency of the stimulus is critical for epilepsy sufferers: the spatial frequency found to elicit peak levels of epileptiform activity is around 2.5 (Wilkins et al., 1975) and 4 cycles/degree (Wilkins et al., 1979). This is approximately the frequency where one cycle would span one hypercolumn.

Excitation alone does not result in synchronised activity: inhibition is critical to synchronisation (Fries, 2009). The length of the refractory period, the time it takes for a cell to recover after firing, is important to synchronise activity. The time taken by the refractory period can be influenced by inhibition from other cells. Extracellular inhibition can make the refractory period of the cell slightly longer, and therefore change its phase slightly. In this manner, the responses of cells can be shifted closer together in time every time they fire, until they fire at the same time. The neurotransmitter GABA is thought to have a role in controlling this phase locking (Tiesinga and Sejnowski, 2009), and since GABA levels are reduced in migraine (Brighina et al., 2009), this could result in differences in synchronisation in migraine populations.

Synchronisation in Clinical Populations

There is evidence to suggest that synchronisation abnormalities affect clinical populations: Traub, Whittington, Buhl, LeBaue, Bibbig, Boyd, Cross, and Baldeweg (2001) provided evidence to suggest that oscillatory behaviour might be associated with epilepsy seizure. They found that bursts of very fast synchronised oscillations overlay epileptiform activity in child sufferers. Using an animal model they also found in vitro that the drug carbachol caused more sustained gamma oscillations, rather than bursts of fast oscillations, and also reduces the chances of epileptogenic activity. From this they conclude that lack of GABA controlled gamma oscillations result in this fast bursting, and
Oscillatory activity has associations with visual abnormalities in the general population, as well as epilepsy sufferers. Adjamian, Holliday, Barnes, Hillebrand, Hadjipapas, and Singh (2004) showed that illusions reported by non-clinical populations correlated with abnormal neural oscillations. In particular, gamma oscillations correlated with illusions and discomfort ratings, and there was a peak in these oscillations at three cycles/degree. This is important as three cycles/degree gratings have been shown to cause both epileptogenic activity in epilepsy sufferers (e.g. Wilkins et al., 1979), and also highest discomfort judgements in non-clinical populations (Wilkins et al., 1984). This suggests that illusions, epileptoform activity and 3 cycles/degree stimuli might be associated with abnormalities in gamma oscillations.

It is possible that abnormalities in the gamma band are one potential cause of migraine. These oscillations might be problematic if the activity in migraineurs is oversynchronised in the interictal period. It has been suggested that the migraine sufferer usually exhibits diffuse oversynchronisation on presentation of visual stimuli during the interictal period (Angelini, Tommaso, Guido, Hu, Ivanov, Marinazzo, Nardulli, Nitti, Pellicoro, Pierro, and Stramaglia, 2004). More specifically, oversynchronisation was found in the alpha band (in response to 9-24Hz stimulation) of interictal MO sufferers on viewing flashing visual stimuli (de Tomasso, Marinazzo, Guido, Libro, Stramaglia, Nitti, Lattanzi, Angelini, and Pellicoro, 2005). This could potentially put a metabolic strain on the visual system, and builds up over the course of the migraine cycle. Finally this strain reaches critical levels and the attack begins. During an attack, this usual state of oversynchronisation might be reversed. One mechanism for reversing the oversynchronisation would be the CSD: the travelling wavefront propagates from the point of origination (Contreras, 2007), and the cortex remains inactive for a period of time, independent of any stimulation (ao, 1944). This would be important in order to reset the activity levels. Speculatively, it has been suggested that a failure of gamma band synchronisation could result in mosaic like patterns (Billock and Tsou, 2004), similar to those reported by some sufferers during an attack (Sacks, 1992).

The degree of inter-cortical synchronisation is affected by a number of properties of the stimulus. For example, Lima, Singer, Chen, and Neuenschwander (2010) showed synchronisation in response to gratings in the monkey V1. Introducing a second grating, thus creating a plaid stimulus, reduces this synchronisation. Plaids are less likely to elicit abnormal neural activity in photosensitive
10.9. SYNCHRONISATION OF NEURAL RESPONSES

epilepsy sufferers than single gratings (Wilkins et al., 1979). Additionally, non-clinical observers consider square-wave gratings to be more uncomfortable to view than plaid patterns (Conlon et al., 1998). Onat, Nortmann, Rekauzke, König, and Jancke (2011b) showed that the spatial frequency of the stimulus affects the propagation of travelling waves of activation; higher grating spatial frequency (0.4cpd) led to less pronounced propagation than lower spatial frequency gratings (around 0.1cpd) in the cat visual cortex. Critically for the current work, Onat, Koenig, and Janke (2011a) demonstrated very different patterns of activity on viewing natural images compared to gratings. When viewing natural scene movies, not all neurons in V1 were active, and those that were showed variable latency of onset of activity. When viewing gratings, the pattern of activation was very different: there were ripple-like patterns of activation through the cortex. Overall, the gratings showed higher activation than the natural scene movies, which provides support that these stimuli result in overactivation of the cortex compared to natural scenes, possible due to sparse coding for the natural scenes. In fact, 32% more contrast was needed for the natural images to produce the same level of activation to match the cortical response to gratings, showing that there might be a substantial saving in energy due to efficient coding for natural images.

In summary, it might be the case that in normal brains, the idling state is characterised by high amplitude, synchronised activity in some frequency bands. Desynchronisation, and thus lower overall activity, occurs on processing sensory stimuli. In the case of visual stimuli, this is usually contained within the alpha band (8-25Hz). This is related to, but distinct from the percept-related gamma activity. Migraineurs might show increased synchronisation of neural activity in the interictal state, including oversynchronisation in response to stimulation, as opposed to desynchronisation. This oversynchronisation might be higher amplitude than in controls, and therefore it might be metabolically unstable. The migraine attack might be the crisis to resolve this constant level of oversynchronised activity. The travelling wave begins, and propagates throughout the cortex. Reduced levels of GABA means that the activity is not restricted, and spreads through large areas of the cortex. This wavefront of activity is phenomenally experienced as the illusory effects of the migraine aura. Although the wavefront might demand much activity, the cortex is inactive after it has passed, therefore overall this is not enough to increase the haemodynamic response. As the wavefront ignores cortical mapping within a hemisphere, and evokes activity from all units at the wavefront, this disrupts the synchronised activity. The breakdown could account for some of the strange visual phenomena reported from the aura. As the normal activity
10.9. SYNCHRONISATION OF NEURAL RESPONSES

spontaneously returns, so does the haemodynamic response, and thus vasodilation begins the headache.

Oversynchronisation could account for some of the functional differences between migraineurs and controls. For example, the lack of reduced contrast sensitivity in migraineurs could be the result of oversynchronisation reducing the effective signal-to-noise ratio. If many units are responding due to synchronisation, these additional units could drown out the signal from those actually responding to the stimulus. This might effectively add noise to the signal. This explanation could also account for the difficulties in detecting orientation, if other units selective for other orientations are also active through oversynchronisation. Oversynchronisation explanations can account for why pop-out visual search tasks are unaffected in migraineurs. As gamma band synchronisation is thought to be responsible for grouping similar, or coherent features, one pop-out feature which is very different from the background/distractors would not evoke the synchronisation in this band. However, in motion coherence tasks the oversynchronisation might lead to excessive binding of both the target and the distractor dots, thus leading to difficulties in discriminating the direction of global motion.

This account of migraine is still in its very early stages, and therefore mostly untested. Controlled experiments are needed to ascertain whether this theory holds. Not all evidence is in agreement: it has been suggested that widespread synchronised activity is a feature of epilepsy, but not migraine (Haigh, Karanovic, Wilkinson, and Wilkins, 2012). In this study, Haigh et al. (2012) used either static, drifting or vibrating gratings. It is thought that vibrating gratings will increase neural synchrony, whilst drifting gratings will decrease synchrony. Previous work in epilepsy sufferers showed increased paroxysmal responses to vibrating, but not drifting gratings (Binnie et al., 1985), which is indicative that epilepsy sufferers showed increased synchrony for the problematic gratings. By contrast, Haigh et al. (2012) found that migraineurs report more aversion on viewing either drifting or vibrating gratings compared to static ones. This is not indicative of a similar process of increased synchrony as found in epilepsy sufferers.

Also, the explanation involving synchronisation is limited in that it really only explains the MA disorder. Although the functional and EEG results are similar for MO, some differences have been found. Therefore it might be the case that these are two theoretically distinct disorders, despite appearing similar. It is possible that MO are actually MA, but that the aura is a different
sensory modality, and therefore not as easily noticeable as visual symptoms. It could also be the case that MA are simply the more extreme version of the same underlying disorder - it has been suggested that MO might simply have a ‘silent aura’, as it is possible that they simply do not reach a threshold for the illusions, but still suffer from the rest of the symptoms of migraine (see Purdy, 2008, for a discussion). As this is yet uncertain, caution must be excised before applying this theory to both.

10.10 Conclusion

Migraine is a complicated, multifaceted disorder with large individual variation. Evidence seems to suggest that this might be a term encompassing many subtly different disorders, but one important distinction is between MA and MO sufferers. Migraine aura is strongly associated with vision in particular, and visual processing abnormalities have been shown in MA and MO sufferers compared to controls.

Previous, well established explanations of migraine have included hyperexcitation, reduced inhibition, reduced habituation and excessive noise in the cortex. These explanations might not be mutually exclusive. However, none of these explanations can completely account for the EEG and visual processing abnormalities seen in migraineurs. There is a new explanation emerging based on abnormal synchronisation in MA sufferers. This has the advantage of providing an explanation for the elevated EEG activity reported in migraineurs, and also much of their psychophysical performance. It can also explain the origins of the CSD, and provide an account of the mechanisms on the level of populations of cells, thus incorporating neurotransmitter imbalances. The theory has not been thoroughly investigated yet in a controlled environment, however preliminary investigations provide evidence to support the hypothesis of oversynchronisation in the migraine brain, especially in the alpha and gamma bands. These bands are thought to be responsible for the processing of visual stimuli, and for the binding of coherent features respectively.

Despite much disagreement in the literature, there is some unanimity on some aspects. For example, migraine is affected by visual stimulation, and this is critical for the current work. Excess activity is thought to be the consequence of viewing periodic stimuli such as gratings and flashing lights. The oversynchronisation explanation can account for why these stimuli might be problematic for migraine sufferers - gratings produce ripple-like oscillations.
in the brain, whereas natural stimuli evoke much less activity. A tendency to oversynchronise visual responses might have the effect of exacerbating the response to gratings further, thus non-clinical populations might find them uncomfortable, but in migraineurs this might elicit a migraine attack. Future research might be well directed in investigating this potential explanation, possibly using more complex stimuli that will evoke synchronised responses in the different bands, and see how synchronisation of responses changes over the course of the migraine cycle of the MA sufferer. It is still unclear if there is a functional distinction between MA and MO, however in terms of theoretical explanations, it might be worth discriminating between the two in future investigations.
CHAPTER 11

CONCLUSION

11.1 Visual Discomfort

Visual discomfort is the subjective unpleasant sensations, such as headache, eyestrain and distortions of vision, reported on viewing particular stimuli such as striped patterns, blurred stimuli and those creating cue conflict. There seems to be a wide variety of types of discomfort, such as those creating headache compared to that resulting in dry eyes (Sheedy et al., 2003). It is likely that there are a variety of causes for the differing symptoms. For example, cue conflict between accommodation and vergence might cause headaches in some individuals (e.g. Shibata et al., 2011), whereas reading low contrast text might cause more eyestrain reports (Nahar et al., 2007). In this thesis I have focused on investigating two possible theoretical causes of visual discomfort: accommodation and neural coding efficiency. It must be emphasised that this is not an exhaustive account of all potential sources of discomfort from visual stimuli. There are many of other stimulus properties that can also cause discomfort, such as colour (Haigh, Allen, and Wilkins, 2011). These attributes can also be explained with the main argument that uncomfortable stimuli are difficult stimuli for the visual system to process. The work in this thesis was an investigation of two possible ways in which some stimuli could pose a problem for the visual system.

11.2 Neural Coding Efficiency

Previous work into visual discomfort from static stimuli showed there to be effects of spatial frequency manipulations on discomfort judgements. This
was both in striped gratings (Wilkins et al., 1984) and filtered noise stimuli (Fernandez and Wilkins, 2008). Images with different amplitude spectra to those typical of natural images were found to increase discomfort judgements. In particular, Chapter 2 showed that filtered noise stimuli with an increase in amplitude at midrange spatial frequencies (relative to natural images) were judged to be more uncomfortable than those with the statistics of natural images. This has been shown not to be purely an effect of either physical (Fernandez and Wilkins, 2008) or perceived contrast (see Chapter 2). The effect could be accounted for using inefficient neural coding. Sparse coding models have been suggested to account for the optimisation of the visual system to natural images (e.g. Field, 1994). Images with the statistics typical of natural images will be coded efficiently by the visual system. This is important in terms of conserving metabolic energy (Lennie, 2003; Barlow, 1961). Conversely, images with very different statistics compared to natural images will not be able to be coded efficiently. These stimuli might therefore result in metabolically inefficient, excessive responses. This might result in discomfort.

Evidence for this possibility was investigated in Chapter 3. The sparse coding model in the work by Field (1994) suggested that natural images might efficiently be encoded by progressive filtering through the stages of the visual system. High response kurtosis (with a mode response of zero), was taken to be a measure of sparseness. This was shown to increase with each progressive stage of filtering for all images tested. Some of the uncomfortable images were found not to produce sparse responses with this model - instead, the responses were more broadly distributed, meaning many neurons being strongly active. This would support the notion that increased metabolic expenditure from excessive activity might provide a source of discomfort.

Clinical populations such as migraine and epilepsy sufferers might suffer from excessive responses in general (e.g. Wilkins et al., 1979; Aurora and Wilkinson, 2007). If stimuli do not allow themselves to be coded efficiently as they lack the statistics characteristic of natural images, then these might exacerbate the problem for these populations. This could result in a more extreme cortical response, and therefore increased discomfort judgements. Additionally, striped patterns that have also been shown to elicit discomfort responses in non-clinical populations (Wilkins et al., 1984) have also been shown to drive excessive responses in clinical populations that could also potentially precipitate seizures (Radhakrishnan et al., 2005) and migraine attacks (Mulleners et al., 2001) in those susceptible.
11.3. INADEQUATE ACCOMMODATION CUES

In the case of migraine, there has already been much research into the perceptual differences of sufferers compared to the general population. Attempts have been made to characterise and explain the origins of the disorder. Many potential explanations for the overresponsiveness have been suggested, including excessive neural noise (e.g. Wagner et al., 2010), and hyperexcitability (Aurora and Wilkinson, 2007). It could also be the case that disorders of synchronisation of responses could lead to the problem. Oversynchronisation interictally might manifest itself as desynchronisation ictally, and thus precipitate the attack. This possibility could be examined psychophysically, but this is left for future research.

However, results of the model in Chapter 3 showed that not all classes of uncomfortable images were found to show the same response distribution. For example, filtered noise images showed response distributions quite similar to those of natural images. By contrast, other striped patterns tended to show a broader distribution of responses. Therefore it could be the case that there is an alternative explanation for the cause of discomfort from noise patterns compared to discomfort from striped patterns: it does not seem to be the case that discomfort from noise patterns is due to response distribution. One possibility is accommodation responses.

11.3 Inadequate Accommodation Cues

Discomfort might arise from the stimulus providing inadequate cues to accommodation. The accommodative system is thought to operate to maximise retinal image contrast, and this requires feedback from the stimulus. The stimulus to accommodation needs to contain contrast at certain spatial frequencies in order to achieve this (e.g. Charman and Tucker, 1977; Charman, 1979). It has been argued that this needs to be in the form of maximum contrast gradients (Day et al., 2009a).

The results of Chapter 4 showed that the loss of information at high spatial frequencies increased discomfort judgements both for simple and natural stimuli. Results of an optical model of the change in contrast from the incident to the retinal image showed that this could be due to a smaller change in contrast gradient from the less comfortable stimuli compared to the more comfortable stimuli. Little change in contrast gradient with defocus could be the cause of the discomfort as there is an insufficient cue to accommodation responses.
Chapter 4 showed the same result that increasing blur resulted in increased discomfort judgements. This was irrespective of whether the blurring function was simply a removal of harmonics from simple waveforms, or Gaussian or sinc filtering of natural images. The results of Chapter 5 show that it does make a difference where the blur is located in the image. The global blurring of Chapter 5 was at the focal point of the image, and could therefore affect accommodation responses. By contrast, the blurring in Chapter 4 introduced due to depth-of-field effects, did not affect discomfort judgements. This was either alone, or when providing another depth cue that could potentially influence a pre-existing accommodation-vergence conflict typical of stereoscopic viewing. The null effect of DOF blur was present even though the DOF blur gradient added were much larger than should be there in the scene, assuming human optics. Therefore, it is concluded that DOF, which does not affect accommodation responses directly, does not influence discomfort. However, blurring the fixation point, which can affect accommodative responses, does influence discomfort. This strengthens the argument that accommodative responses might be the cause of the discomfort in these cases, not just blurring per se.

One of the unanswered questions here is whether the potential accommodative difficulties are purely due to fatigue in the motor response, due to increased microfluctuations, or discrepancies in the accommodation control system. As with inefficient neural coding, it could be the case that uncertainty in the accommodation control systems cause discomfort due to inefficient processing: constantly reanalysing the accommodative cue and attempting to send signals to the muscles to bring it into focus. This could be metabolically costly in terms of processing efficiency - a certain response would perhaps lead to calculating this once only, instead of multiple recalibrations. Alternatively, the uncertainty in the response could result in increased microfluctuations (Day et al., 2009a). These could result in simple muscle fatigue, which could cause the discomfort.

It would be interesting to assess whether the discomfort is a case of muscle fatigue or sensory discrepancy. This could either be done using an autorefractor, to correlate discomfort and microfluctuation amplitude. Alternatively, a pinhole experiment might give a more causal result. When observers view a stimulus through a pinhole aperture, it has been reported that microfluctuations can be reduced (Miles, Judge, and Optican, 1987; Morrison et al., 2010). Therefore, if uncomfortable stimuli are viewed though a pinhole aperture, and they are no longer uncomfortable, then it might be due to the stopping of
11.4 Impact of Visual Discomfort

The impact of uncomfortable stimuli on performance was also measured, to assess how disruptive these stimuli might be to general populations. A visual search task in Chapter 6 showed no evidence that either filtered noise or striped backgrounds were detrimental to performance. This was either in terms of reaction time or accuracy. There was evidence that the effects on performance shown could have been due to surround suppression, but no evidence of discomfort. Performance on a reading task in Chapter 7 also showed little impact of discomfort. There were no effects of text manipulation on reading performance in a naturalistic reading task. However, it is possible that differences in uncontrolled variables such as familiarity of words and differences in linguistic ability might have masked results. Reading was faster with two eyes in this experiment, which is not supportive that the text is difficult to read due to binocular co-ordination issues.

In a more simplified controlled task, differences in reading rate were found as a result of text stripiness. However, this was a complex interaction effect between font type and text stripiness as determined by autocorrelation function. The sans serif ‘Arial’ font was read more quickly with more stripy words, the ‘Times’ font more quickly with the less stripy words. It is unclear why this might be the case - neither coding efficiency nor binocular mismatching explanations can account for this result.

The lack of findings on these tasks could be because there are few individuals who report high levels of discomfort, according to the Conlon et al. (1999) VDS. A wider online survey of individuals, who were not exclusively students, confirmed this finding. It is curious that this does not match the reports from previous work, such as Conlon et al. (1999, 2001); Conlon and Humphreys (2001); Chase et al. (2009); Tosha et al. (2009), however, this was exactly the same questionnaire. I have no explanation for this, other than potential cultural differences in the admission of discomfort. Conlon et al. (1999) is based in Australia, Chase et al. (2009); Tosha et al. (2009) in America. It is possible that the online sample were predominantly British, and therefore there might be cultural differences. However, this is pure speculation based
11.5. THE ROLE OF CONTRAST

on stereotyping. There is some evidence that British males have higher tolerance to heat pain compared to South Asians, but there is no evidence to directly compare British and Australian pain thresholds (Watson, Latif, and Rowbotham, 2005). Secondly, there are no questions in the survey regarding nationality. By contrast, the Pattern Glare Test did show results comparable with the rates reported by Evans and Stevenson (2008). This could be as there the questions are directly linked to the stimulus presented at that time, thus removing the memory component.

11.5 The Role of Contrast

Although stimuli argued to be uncomfortable do not directly impact performance in tasks, it could be the case that the stimuli were not strong enough to induce sufficient levels of discomfort. High contrast stimuli could be more uncomfortable. New technological developments are focusing on making displays that are legible under daylight conditions, and one method is to increase overall luminance. Increasing overall luminance also has the effect of creating a wider range of contrast, thus stimuli can be presented at a higher contrast range than before. The results of Chapter 8 suggest that gain control mechanisms might deal with the potential issue of an overall increase luminance itself. However, there were effects of contrast range on discomfort judgements. Higher contrast increased comfort of natural images and text, possibly by increasing clarity. However, higher contrast of uncomfortable stimuli, such as striped gratings, decreased comfort judgements despite the higher clarity judgements.

This result has implications for displays, that users of these brighter displays should be careful what exactly is presented on the screen. Also, this has theoretical implications, supporting the suggestion that striped patterns might be uncomfortable due to excessive responses in the brain. Increasing the contrast would be expected to increase the discomfort, as the problem is exacerbated. This is what was shown in Chapter 8.

11.6 Conclusion

There is evidence to suggest two potential causes that could account for some types of visual discomfort in the general population. Inefficient neural coding could account for the findings from striped gratings, and accommodation responses might be the cause of discomfort from blurred stimuli. It appears to be the case that difficult stimuli for the visual system to process are un-
comfortable. This is possibly due to increased effort required to process them. Individual differences could be due to the nature of the owner’s visual system. For example, those with generally higher levels of cortical noise might find striped patterns more uncomfortable than those with less noise. Those with good optics might be less tolerant than others of blur. Discomfort might tell us something about brain function, but does not present an in escapable and dramatic cause for concern in everyday life for the general population.
25 January 2010

Ethics Reference No: PS0981

Project Title: Visual discomfort and the amplitude spectrum of images

Researchers Name(s): Louise O’Hare

Supervisor(s): Dr Paul Hibbard

Thank you for submitting your application which was considered at the School Ethics Committee meeting on the 22nd January 2010. The following documents were reviewed:

1. Ethical Application Form 22/01/2010
2. Participant Information Sheet 22/01/2010
3. Consent Form 22/01/2010
4. Debriefing Form 22/01/2010
5. Advertisement 22/01/2010

The University Teaching and Research Ethics Committee (UTREC) approves this study from an ethical point of view. Please note that where approval is given by a School Ethics Committee that committee is part of UTREC and is delegated to act for UTREC.

Approval is given for three years. Projects, which have not commenced within two years of original approval, must be re-submitted to your School Ethics Committee.

You must inform your School Ethics Committee when the research has been completed. If you are unable to complete your research within the three year validation period, you will be required to write to your School Ethics Committee and to UTREC (where approval was given by UTREC) to request an extension or you will need to re-apply.

Any serious adverse events or significant change which occurs in connection with this study and/or which may alter its ethical consideration, must be reported immediately to the School Ethics Committee, and an Ethical Amendment Form submitted where appropriate.

Approval is given on the understanding that the ‘Guidelines for Ethical Research Practice’ (http://www.st-andrews.ac.uk/media/UTREC/guidelines%20Feb%202008.pdf) are adhered to.

Yours sincerely

Convenor of the School Ethics Committee

Dr Paul Hibbard (Supervisor)

School Ethics Committee

OR

Convenor of UTREC

UTREC Convenor, Mearfield, 2 St Mary’s Place, St Andrews, KY16 9LY
Email: persethics@st-andrews.ac.uk Tel: 01334 462157
The University of St Andrews is a charity registered in Scotland: No S2014322
Thank you for submitting your application which was considered at the Psychology School Ethics Committee meeting on the 7th April 2010. The following documents were reviewed:

1. Ethical Amendment Form 07/04/2010
2. Participant Information Sheet 07/04/2010
3. Consent Form 07/04/2010
4. Debriefing Form 07/04/2010
5. Advertisement 07/04/2010

The University Teaching and Research Ethics Committee (UTREC) approves this study from an ethical point of view. Please note that where approval is given by a School Ethics Committee that committee is part of UTREC and is delegated to act for UTREC.

Approval is given for three years. Projects, which have not commenced within two years of original approval, must be re-submitted to your School Ethics Committee.

You must inform your School Ethics Committee when the research has been completed. If you are unable to complete your research within the 3 three year validation period, you will be required to write to your School Ethics Committee and to UTREC (where approval was given by UTREC) to request an extension or you will need to re-apply.

Any serious adverse events or significant change which occurs in connection with this study and/or which may alter its ethical consideration, must be reported immediately to the School Ethics Committee, and an Ethical Amendment Form submitted where appropriate.

Approval is given on the understanding that the "Guidelines for Ethical Research Practice" (http://www.st-andrews.ac.uk/media/UTREC/guidelines620Feb%2008.pdf) are adhered to.

Yours sincerely

Convener of the School Ethics Committee

OR

Convener of UTREC

Dr Paul Hibbard (Supervisor)
School Ethics Committee
31 May 2010

<table>
<thead>
<tr>
<th>Ethics Reference No:</th>
<th>PS6415</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Project Title:</strong></td>
<td>Text characteristics and reading</td>
</tr>
<tr>
<td><strong>Researchers Name(s):</strong></td>
<td>Louise O'Hare, Catherine Gordon</td>
</tr>
<tr>
<td><strong>Supervisor(s):</strong></td>
<td>Dr Paul Hibbard</td>
</tr>
</tbody>
</table>

Thank you for submitting your application which was considered at the Psychology School Ethics Committee meeting on the 31 May 2010. The following documents were reviewed:

1. Ethical Application Form 31/05/2010
2. Participant Information Sheet 31/05/2010
3. Consent Form 31/05/2010
4. Debriefing Form 31/05/2010
5. Advertisement 31/05/2010

The University Teaching and Research Ethics Committee (UTREC) approves this study from an ethical point of view. Please note that where approval is given by a School Ethics Committee that committee is part of UTREC and is delegated to act for UTREC.

Approval is given for three years. Projects, which have not commenced within two years of original approval, must be re-submitted to your School Ethics Committee.

You must inform your School Ethics Committee when the research has been completed. If you are unable to complete your research within the 3 year validation period, you will be required to write to your School Ethics Committee and to UTREC (where approval was given by UTREC) to request an extension or you will need to re-apply.

Any serious adverse events or significant change which occurs in connection with this study and/or which may alter its ethical consideration, must be reported immediately to the School Ethics Committee, and an Ethical Amendment Form submitted where appropriate.

Approval is given on the understanding that the “Guidelines for Ethical Research Practice” (http://www.st-andrews.ac.uk/mcell/UTREC/guidelines/K20Feb52008.pdf) are adhered to.

Yours sincerely

Convenor of the School Ethics Committee

OR

Convenor of UTREC

Ccs: Dr Paul Hibbard (Supervisor)
School Ethics Committee
35 June 2010

<table>
<thead>
<tr>
<th>Ethics Reference No:</th>
<th>PS6415</th>
</tr>
</thead>
<tbody>
<tr>
<td>Project Title:</td>
<td>Text characteristics and reading</td>
</tr>
<tr>
<td>Researchers Name(s):</td>
<td>Louise O'Hare and Catherine Gordon</td>
</tr>
<tr>
<td>Supervisor(s):</td>
<td>Dr Paul Hibbard</td>
</tr>
</tbody>
</table>

Thank you for submitting your application which was considered at the Psychology School Ethics Committee meeting on the 15th June 2010. The following documents were reviewed:

1. Ethical Amendment Form 15/06/2010
2. Participant Information Sheet 15/06/2010
3. Sample of Test Pattern 15/06/2010

The University Teaching and Research Ethics Committee (UTREC) approves this study from an ethical point of view. Please note that where approval is given by a School Ethics Committee that committee is part of UTREC and is delegated to act for UTREC.

Approval is given for three years. Projects, which have not commenced within two years of original approval, must be re-submitted to your School Ethics Committee.

You must inform your School Ethics Committee when the research has been completed. If you are unable to complete your research within the 3 three year validation period, you will be required to write to your School Ethics Committee and to UTREC (where approval was given by UTREC) to request an extension or you will need to re-apply.

Any serious adverse events or significant change which occurs in connection with this study and/or which may alter its ethical consideration, must be reported immediately to the School Ethics Committee, and an Ethical Amendment Form submitted where appropriate.

Approval is given on the understanding that the 'Guidelines for Ethical Research Practice' (http://www.st-andrews.ac.uk/media/UTREC/unit/UTREC_guidelines%202008.pdf) are adhered to.

Yours sincerely

Convenor of the School Ethics Committee

Cos Dr Paul Hibbard (Supervisor)
School Ethics Committee

Convenor of UTREC

UTREC Convenor, Mansfield, 2 St Mary's Place, St Andrews, KY16 9LY
Email: jayethico@st-andrews.ac.uk Tel: 01334 662157
The University of St Andrews is a charity registered in Scotland: No SCO12532

274
5 April 2011

<table>
<thead>
<tr>
<th>Ethics Reference No:</th>
<th>PS6415 (Amendment)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Project Title:</td>
<td>Text characteristics and reading</td>
</tr>
<tr>
<td>Researchers Name(s):</td>
<td>Louise O’Hare</td>
</tr>
<tr>
<td>Supervisor(s):</td>
<td>Dr Paul Hibbard</td>
</tr>
</tbody>
</table>

Thank you for submitting your application which was considered at the Psychology School Ethics Committee meeting on the 4th April 2011. The following documents were reviewed:

1. Ethical Amendment Form

The University Teaching and Research Ethics Committee (UTREC) Approves this study from an ethical point of view. Please note that where approval is given by a School Ethics Committee that committee is part of UTREC and is delegated to act for UTREC.

Approval is given for three years. Projects, which have not commenced within two years of original approval, must be re-submitted to your School Ethics Committee.

You must inform your School Ethics Committee when the research has been completed. If you are unable to complete your research within the 3 three year validation period, you will be required to write to your School Ethics Committee and to UTREC (where approval was given by UTREC) to request an extension or you will need to re-apply.

Any serious adverse events or significant change which occurs in connection with this study and/or which may alter its ethical consideration, must be reported immediately to the School Ethics Committee, and an Ethical Amendment Form submitted where appropriate.

Approval is given on the understanding that the ‘Guidelines for Ethical Research Practice’ (http://www.st-andrews.ac.uk/media/UTREC-guidelines0524Feb022008.pdf) are adhered to.

Yours sincerely

Convenor of the School Ethics Committee

Cc Dr Paul Hibbard (Supervisor)
School Ethics Committee

UTREC Convener, Mansfield, 3 St Mary’s Place, St Andrews, KY16 9UY
Email: utrec@st-andrews.ac.uk Tel: 01334 642866
The University of St Andrews is a charity registered in Scotland: No SC013532
University of St Andrews

University Teaching and Research Ethics Committee

26 October 2010

Ethics Reference No: PS6914

Project Title: Effect of blur on visual discomfort

Researchers Name(s): Louise O’Hare

Supervisor(s): Dr Paul Hibbard

Thank you for submitting your application which was considered at the Psychology School Ethics Committee meeting on the 20th October 2010. The following documents were reviewed:

1. Ethical Application Form 21/10/2010
2. Participant Information Sheet 21/10/2010
3. Consent Form 21/10/2010
4. Debriefing Form 21/10/2010
5. Advertisement 21/10/2010
6. Questionnaires 21/10/2010
7. Pattern Sensitivity Test Card 21/10/2010

The University Teaching and Research Ethics Committee (UTREC) approves this study from an ethical point of view. Please note that where approval is given by a School Ethics Committee that committee is part of UTREC and is delegated to act for UTREC.

Approval is given for three years. Projects which have not commenced within two years of original approval, must be re-submitted to your School Ethics Committee.

You must inform your School Ethics Committee when the research has been completed. If you are unable to complete your research within the 3 year validation period, you will be required to write to your School Ethics Committee and to UTREC (where approval was given by UTREC) to request an extension or you will need to re-apply.

Any serious adverse events or significant change which occurs in connection with this study and/or which may alter its ethical consideration, must be reported immediately to the School Ethics Committee, and an Ethical Amendment Form submitted where appropriate.

Approval is given on the understanding that the ‘Guidelines for Ethical Research Practice’ (http://www.st-andrews.ac.uk/media/UTREC/guidelines%20Feb%2009.pdf) are adhered to.

Yours sincerely

Convenor of the School Ethics Committee

OR

Convenor of UTREC

Cns Dr Paul Hibbard (Supervisor)
School Ethics Committee

UTREC Convenor, Mansfield, 3 St Mary’s Place, St Andrews, KY16 9UI
Email: utrec@st-andrews.ac.uk Tel: 01334 462566
The University of St Andrews is a charity registered in Scotland. No SC011253
28 October 2010

<table>
<thead>
<tr>
<th>Ethics Reference No:</th>
<th>PS6914 (Amendment)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Project Title:</td>
<td>Effect of blur on visual discomfort</td>
</tr>
<tr>
<td>Researchers Name(s):</td>
<td>Louise O'Hare</td>
</tr>
<tr>
<td>Supervisor(s):</td>
<td>Dr Paul Hibbard</td>
</tr>
</tbody>
</table>

Thank you for submitting your application which was considered at the Psychology School Ethics Committee meeting on the 29th October 2010. The following documents were reviewed:

1. Ethical Amendment Form 29/10/2010
2. Participant Information Sheet 29/10/2010

The University Teaching and Research Ethics Committee (UTREC) approves this study from an ethical point of view. Please note that where approval is given by a School Ethics Committee that committee is part of UTREC and is delegated to act for UTREC.

Approval is given for three years. Projects, which have not commenced within two years of original approval, must be re-submitted to your School Ethics Committee.

You must inform your School Ethics Committee when the research has been completed. If you are unable to complete your research within the 3 three year validation period, you will be required to write to your School Ethics Committee and to UTREC (where approval was given by UTREC) to request an extension or you will need to re-apply.

Any serious adverse events or significant change which occurs in connection with this study and/or which may alter its ethical consideration, must be reported immediately to the School Ethics Committee, and an Ethical Amendment Form submitted where appropriate.

Approval is given on the understanding that the ‘Guidelines for Ethical Research Practice’ (http://www.st-andrews.ac.uk/media/UTREC/guidelines%20Feb%2008.pdf) are adhered to.

Yours sincerely

Convenor of the School Ethics Committee

Ces Dr Paul Hibbard (Supervisor)
School Ethics Committee

OR

Convener of UTREC
5 April 2011

<table>
<thead>
<tr>
<th>Ethics Reference No.</th>
<th>PS7130 (Amendment)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Project Title:</td>
<td>The effects of brightness in high-brightness displays on visual discomfort</td>
</tr>
<tr>
<td>Researchers Name(s):</td>
<td>Louise O'Hare, Dr Paul Hibbard, Caroline Foubister</td>
</tr>
<tr>
<td>Supervisor(s):</td>
<td>Dr Paul Hibbard</td>
</tr>
</tbody>
</table>

Thank you for submitting your application, which was considered at the Psychology School Ethics Committee meeting on the 4th April 2011. The following documents were reviewed:

1. Ethical Amendment Form

The University Teaching and Research Ethics Committee (UTREC) approves this study from an ethical point of view. Please note that where approval is given by a School Ethics Committee, that committee is part of UTREC and is delegated to act for UTREC.

Approval is given for three years. Projects, which have not commenced within two years of original approval, must be re-submitted to your School Ethics Committee.

You must inform your School Ethics Committee when the research has been completed. If you are unable to complete your research within the three-year validation period, you will be required to write to your School Ethics Committee and to UTREC (where approval was given by UTREC) to request an extension or you will need to re-apply.

Any serious adverse events or significant change which occurs in connection with this study and/or which may alter its ethical consideration, must be reported immediately to the School Ethics Committee, and an Ethical Amendment Form submitted where appropriate.

Approval is given on the understanding that the 'Guidelines for Ethical Research Practice' (http://www.st-andrews.ac.uk/media/UTREC/guidelines32020e86208.pdf) are adhered to.

Yours sincerely

Convenor of the School Ethics Committee

[Signature]

Dr Paul Hibbard (Supervisor)
School Ethics Committee

UTREC Convenor, Marshfield, 3 St Mary's Place, St Andrews, KY16 9UY
Email: untrc@st-andrews.ac.uk Tel: 0334-462866
The University of St Andrews is a charity registered in Scotland: No SC015232
University of St Andrews

University Teaching and Research Ethics Committee

9 December 2010

<table>
<thead>
<tr>
<th>Ethics Reference No:</th>
<th>PS7078</th>
</tr>
</thead>
<tbody>
<tr>
<td>Project Title:</td>
<td>Visual discomfort survey</td>
</tr>
<tr>
<td>Researchers Name(s):</td>
<td>Louise O'Hare</td>
</tr>
<tr>
<td>Supervisor(s):</td>
<td>Dr Paul Hibbard</td>
</tr>
</tbody>
</table>

Thank you for submitting your application which was considered at the Psychology School Ethics Committee meeting on the 8th December 2010. The following documents were reviewed:

1. Ethical Application Form  08/12/2010
2. Participant Information Sheet  08/12/2010
3. Consent Form  08/12/2010
4. Debriefing Form  08/12/2010
5. Questionnaire  08/12/2010

The University Teaching and Research Ethics Committee (UTREC) approves this study from an ethical point of view. Please note that where approval is given by a School Ethics Committee that committee is part of UTREC and is delegated to act for UTREC.

Approval is given for three years. Projects, which have not commenced within two years of original approval, must be re-submitted to your School Ethics Committee.

You must inform your School Ethics Committee when the research has been completed. If you are unable to complete your research within the 3 three year validation period, you will be required to write to your School Ethics Committee and to UTREC (where approval was given by UTREC) to request an extension or you will need to re-apply.

Any serious adverse events or significant change which occurs in connection with this study and/or which may alter its ethical consideration, must be reported immediately to the School Ethics Committee, and an Ethical Amendment Form submitted where appropriate.

Approval is given on the understanding that the 'Guidelines for Ethical Research Practice' (http://www.st-andrews.ac.uk/medialibrary/UTREC/guidelines%20Feb%202008.pdf) are adhered to.

Yours sincerely

Convener of the School Ethics Committee

Dr Paul Hibbard (Supervisor)
School Ethics Committee

OR

Convener of UTREC
Project Title: Visual Discomfort Survey

Researchers Name: Louise O'Hare

Supervisor: Dr Paul Hibbard

Department/Unit: School of Psychology

Ethical Approval Code: FS7078

(Attribution of original application)

Original Application Approval Date: 7 March 2011

Amendment Application Approval Date: 10 October 2011

Ethical Amendment Approval

Thank you for submitting your amendment application which was considered at the Psychology School Ethics Committee meeting on the 5th October 2011. The following documents were reviewed:

1. Ethical Amendment Application Form 10/10/2011
2. Participant Information Sheet 10/10/2011
3. Consent Form 10/10/2011
4. Debriefing Form 10/10/2011
5. Questionnaires 10/10/2011

The University Teaching and Research Ethics Committee (UTREC) approves this study from an ethical point of view. Please note that where approval is given by a School Ethics Committee that committee is part of UTREC and is delegated to act for UTREC.

Approval is given for three years from the original application only. Ethical Amendments do not extend this period but give permission to a further amendment to the original approval for research proposal only. If you are unable to complete your research within the original 3 three year validation period, you will be required to write to your School Ethics Committee and to UTREC (where approval was given by UTREC) to request an extension or you will need to re-apply. You must inform your School Ethics Committee when the research has been completed.

Any serious adverse events or significant changes which occur in connection with this study and/or which may alter its ethical consideration, must be reported immediately to the School Ethics Committee, and an Ethical Amendment Form submitted where appropriate.

Approval is given on the understanding that the ‘Guidelines for Ethical Research Practice’ (http://www.st-andrews.ac.uk/media/UTREC/guidelines%207Feb%202018.pdf) are adhered to.

Yours sincerely

Convener of the School Ethics Committee

CoS Dr P. Hibbard (Supervisor) School Ethics Committee
7 March 2011

<table>
<thead>
<tr>
<th>Ethics Reference No:</th>
<th>PS7078 (Amendment)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Project Title:</td>
<td>Visual discomfort survey</td>
</tr>
<tr>
<td>Researchers Name(s):</td>
<td>Louise O'Hare</td>
</tr>
<tr>
<td>Supervisor(s):</td>
<td>Dr. P. B. Hibbard</td>
</tr>
</tbody>
</table>

Thank you for submitting your application which was considered at the Psychology School Ethics Committee meeting on the 7th March 2011. The following documents were reviewed:

1. Ethical Amendment Form 07/03/2011

The University Teaching and Research Ethics Committee (UTREC) approves this study from an ethical point of view. Please note that where approval is given by a School Ethics Committee that committee is part of UTREC and is delegated to act for UTREC.

Approval is given for three years. Projects, which have not commenced within two years of original approval, must be re-submitted to your School Ethics Committee.

You must inform your School Ethics Committee when the research has been completed. If you are unable to complete your research within the 3 three year validation period, you will be required to write to your School Ethics Committee and to UTREC (where approval was given by UTREC) to request an extension or you will need to re-apply.

Any serious adverse events or significant change which occurs in connection with this study and/or which may alter its ethical consideration, must be reported immediately to the School Ethics Committee, and an Ethical Amendment Form submitted where appropriate.

Approval is given on the understanding that the 'Guidelines for Ethical Research Practice' (http://www.st-andrews.ac.uk/med/UTREC/ethicalresearchrectification.pdf) are adhered to.

Yours sincerely,

Convenor of the School Ethics Committee

Cc: Dr. P. B. Hibbard (Supervisor)
School Ethics Committee
19 June 2011

<table>
<thead>
<tr>
<th>Ethics Reference No:</th>
<th>PS7633</th>
</tr>
</thead>
<tbody>
<tr>
<td>Project Title:</td>
<td>Visual search against abstract backgrounds</td>
</tr>
<tr>
<td>Researchers Name:</td>
<td>Louise O'Hare</td>
</tr>
<tr>
<td>Supervisor:</td>
<td>Dr P. B. Hibbard</td>
</tr>
</tbody>
</table>

Thank you for submitting your application which was considered at the Psychology School Ethics Committee meeting on the 8th June 2011. The following documents were reviewed:

1. Ethical Application Form                      08/06/2011
2. Participant Information Sheet                08/06/2011
3. Consent Form                                 08/06/2011
4. Debriefing Form                              10/06/2011

The University Teaching and Research Ethics Committee (UTREC) approves this study from an ethical point of view. Please note that where approval is given by a School Ethics Committee that committee is part of UTREC and is delegated to act for UTREC.

Approval is given for three years. Projects, which have not commenced within two years of original approval, must be re-submitted to your School Ethics Committee.

You must inform your School Ethics Committee when the research has been completed. If you are unable to complete your research within the 3 year validation period, you will be required to write to your School Ethics Committee and to UTREC (where approval was given by UTREC) to request an extension or you will need to re-apply.

Any serious adverse events or significant change which occurs in connection with this study and/or which may alter its ethical consideration, must be reported immediately to the School Ethics Committee, and an Ethical Amendment Form submitted where appropriate.

Approval is given on the understanding that the ‘Guidelines for Ethical Research Practice’ (http://www.st-andrews.ac.uk/media/UTREC/guidelines%20Feb%2008.pdf) are adhered to.

Yours sincerely

Convener of the School Ethics Committee

Cos: Dr P. B. Hibbard (Supervisor)
School Ethics Committee

UTREC Convener, Mazerfield, 3 St Mary's Place, St Andrews, KY16 9DY
Email: esbs@st-andrews.ac.uk Tel: 01334 462866
The University of St Andrews is a charity registered in Scotland: No SC011232
Project Title: Visual Discomfort Survey
Researcher's Name: Louise O'Hare
Supervisor: Dr Paul Hibbard
Department/Unit: School of Psychology
Ethical Approval Code: PS7078
Original Application Approval Date: 8 December 2010
Amendment Application Approval: 22 December 2011

Ethical Amendment Approval

Thank you for submitting your amendment application which was considered at the Psychology School Ethics Committee meeting on the 22 December 2011. The following documents were reviewed:

1. Ethical Amendment Application Form 22/12/2011

The University Teaching and Research Ethics Committee (UTREC) approves this study from an ethical point of view. Please note that where approval is given by a School Ethics Committee that committee is part of UTREC and is delegated to act for UTREC.

Approval is given for three years from the original application only. Ethical Amendments do not extend this period but give permission to an amendment to the original approval research proposal only. If you are unable to complete your research within the original 3 three year validation period, you will be required to write to your School Ethics Committee and to UTREC (where approval was given by UTREC) to request an extension or you will need to re-apply. You must inform your School Ethics Committee when the research has been completed.

Any serious adverse events or significant change which occurs in connection with this study and/or which may alter its ethical consideration, must be reported immediately to the School Ethics Committee, and an Ethical Amendment Form submitted where appropriate.

Approval is given on the understanding that the 'Guidelines for Ethical Research Practice' (http://www.st-andrews.ac.uk/media/UTREC/guidelines%20Feb%202008.pdf) are adhered to.

Yours sincerely

Convenor of the School Ethics Committee

Cos: Dr P. Hibbard (Supervisor)
School Ethics Committee
University of St Andrews
University Teaching and Research Ethics Committee

18 January 2012

<table>
<thead>
<tr>
<th>Ethics Reference No:</th>
<th>PS8321</th>
</tr>
</thead>
<tbody>
<tr>
<td>Project Title:</td>
<td>Depth of Field Blur</td>
</tr>
<tr>
<td>Researchers' Names:</td>
<td>Louise O'Hare, Tingting Zhang, Dr Paul Hibbard, Dr Harold Nefs</td>
</tr>
<tr>
<td>Supervisor:</td>
<td>Dr Paul Hibbard</td>
</tr>
</tbody>
</table>

Thank you for submitting your application which was considered at the Psychology School Ethics Committee meeting on the 17th January 2012. The following documents were reviewed:

1. Ethical Application Form 17/01/2012
2. Advertisement 17/01/2012
3. Participant Information Sheet 17/01/2012
4. Consent Form 17/01/2012
5. Debriefing Form 17/01/2012
6. External Permissions (University of Delft) 17/01/2012

The University Teaching and Research Ethics Committee (UTREC) approves this study from an ethical point of view. Please note that where approval is given by a School Ethics Committee that committee is part of UTREC and is delegated to act for UTREC.

Approval is given for three years. Projects, which have not commenced within two years of original approval, must be re-submitted to your School Ethics Committee.

You must inform your School Ethics Committee when the research has been completed. If you are unable to complete your research within the 3 three year validation period, you will be required to write to your School Ethics Committee and to UTREC (where approval was given by UTREC) to request an extension or you will need to re-apply.

Any serious adverse events or significant change which occurs in connection with this study and/or which may alter its ethical consideration, must be reported immediately to the School Ethics Committee, and an Ethical Amendment Form submitted where appropriate.

Approval is given on the understanding that the 'Guidelines for Ethical Research Practice' (http://www.st-andrews.ac.uk/media/UTREC/guidelines%20Feb%2008.pdf) are adhered to.

Yours sincerely

Convener of the School Ethics Committee

Cos Dr P. Hibbard (Supervisor)
School Ethics Committee

UTREC Convener, Mansfield, 3 St Mary’s Place, St Andrews, KY16 9UY
Email: ethics@students.st-andrews.ac.uk Tel: 01334 462866
The University of St Andrews is a charity registered in Scotland: No SC013532


BIBLIOGRAPHY


BIBLIOGRAPHY


BIBLIOGRAPHY


