Confidence and feedback in visual perceptual learning

Jordi M. Asher

Department of Psychology

University of Essex

This dissertation is submitted for the degree of

Doctor of Philosophy

August 2019

“So Long, and Thanks for All the Fish"

— Douglas Adams

Declaration

I declare that, except where specific reference is made to the work of others, the contents of this thesis is the result of my own work, and has not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university.

Jordi M. Asher
August 2019

Acknowledgements

I would like to first say an enormous thank you to Professor Paul Hibbard who is the Buzz Lightyear of supervisors. Paul, who goes to infinity and beyond for all of his students, deserves the supervisor of a *lifetime* award. Thank you to my collaborators Professor Vincenzo Romei (Chapters 3 & 4) and Dr Louise O'Hare (Chapter 3) and my second supervisor Dr Loes van Dam.

My sincere appreciation goes out to the Department of Psychology in general for making a welcoming, friendly and inclusive environment in which to work in. A special thanks to the department's administration and tech teams for all the support they have provided throughout the years, and to all the people who have sat countless hours in the dark pushing buttons to make this research possible.

I gratefully acknowledge the Scholarship from the Department of Psychology at the University of Essex in allowing me the opportunity to undertake my PhD. Thanks also to the ESSEXLab for the Seecdcorn Funding Grant and to PsyPAG for a research bursary, both of which funded the research undertaken in Chapter 4. Thanks also to the Applied Vision Association and the Experimental Psychology Society for funding towards annual international conferences in order to present my research.

I would also like to say a heartfelt thank you to family and friends who have patiently supported my never-ending workload. Thank you and all the best to those friends with whom I shared countless days with, in our research office. Finally thank you to my examiners Dr Martin Lages and Dr Kevin Dent for your insightful questions and valuable corrections.

Abstract

The classical view that perceptual learning is highly specific to the stimulus properties used during training is being challenged by accumulating evidence of transfer under specific conditions. Perceptual learning in the adult brain may result from an element of cortical reorganisation at a perceptual level or a reorganisation of decision weights at a higher level. This thesis focuses on individual differences in perception, and the mechanisms of perceptual learning. Since the role of feedback is not usually the primary area of interest for perceptual learning research, this was one of the main focus points of the methodology used. Perception and perceptual learning were assessed for a range of tasks at local and global levels, including depth perception in random dot stereograms, and global motion and form coherence. The results of these experiments established that transfer occurred when the task matched the tuning of the global processing area, even when the untrained task was a locally processed task. The results also provided evidence of robust learning in local and global tasks with or without feedback, as long as easy and difficult trials were interleaved within the same task. Furthermore, in some conditions internal feedback provided better learning at sub-threshold stimulus levels than did explicit external feedback. These results suggest that, for low stimulus intensities, external feedback may reduce observers' confidence in their own perceptual decisions. Confidence in perceptual decisions was a key factor throughout all the studies, and where this was measured it was found to be highly correlated with performance. Confidence in perceptual decision and accuracy was low for a depth perception study using anti-correlated random-dot stereogram (ACRDS). The binocular energy model of neural responses predicts that depth from binocular disparity may be perceived in the reversed-direction when the contrast of dots are anti-correlated. Depth was mostly perceived in the correct direction for ACRDS conditions with some inconsistencies. These differences are likely to reflect the inconsistent depth signals, across scale and across first- and second-order channels, elicited by anticorrelated stimuli. To ensure that the rich variability embedded within the individual differences in heterogeneous data sets were not excluded from analyses, mixed effects models were employed throughout the thesis. This technique considers between-individual variation as a random factor which made it possible to investigate behavioural differences between individuals with migraine and control groups. Specifically we evaluated the spatial extent of excitatory and inhibitory interactions using a classic lateral masking task. Overall, contrast thresholds in the baseline condition for the migraine group were lower than those in the control group. There was no difference in the degree of lateral interaction in the migraine group. The results suggest the that impaired performance in perceptual tasks in individuals with migraine may not be as a results of altered local mechanisms.

Table of contents

List of figures 13

List of tables 15

1 Introduction 19

1.1 An introduction to visual perception 19
1.1.1 The historical study of visual perception 20
1.1.2 Visual perception and the brain 22
1.1.3 The modern approach to understanding visual perception 24
1.1.4 Neural plasticity and perception 26
1.2 Perception and Learning 28
1.2.1 Defining perceptual learning 29
1.2.2 Differentiating perceptual learning from other types of learning . . . 31
1.2.3 Measuring Perceptual Learning 34
1.2.4 Perceptual learning in early vision research 34
1.3 The functionally organised visual system 38
1.3.1 Pre-cortical area 39
1.3.2 Sub-cortical areas 39
1.3.3 The cortical hierarchy 42
1.3.4 Feed-forward and feed-back connections 46
1.4 Integration and Segmentation 49
1.4.1 Motion Perception 50
1.4.2 The physiology of motion perception 52
1.4.3 Depth Perception 55
1.4.4 Global Form 59
1.5 Perceptual learning for visual rehabilitation 61
1.5.1 Visual impairments 61

1.5.2 Neuro-modulation and rehabilitation 63
1.5.3 The debate against cortical reorganisation 63
1.6 Mechanisms and Models of Perceptual Learning 66
1.6.1 What, Where Why? Mechanisms of Perceptual Learning 66
1.6.2 How? Models of Perceptual Learning 74
1.7 Feedback in perceptual learning 78
1.8 Chapter Outline 82

2 Inclusive psychophysics: embracing variability in psychophysical performance 85
2.1 Introduction 85
2.2 Psychophysical methods 87
2.3 Fitting the best Sigmoid 92
2.3.1 Range of Levels 93
2.3.2 Spread of Levels 94
2.3.3 Number of Repetitions 95
2.4 Data Collection Methods 95
2.4.1 Method of Constant Stimuli 95
2.4.2 Adaptive Procedures (staircases) 96
2.5 The psychometric function 98
2.5.1 Defining the psychometric function 99
2.5.2 Building the generic psychometric function 101
2.6 The psychometric function and individual differences 102
2.6.1 Inter-individual differences 105
2.6.2 Intra-individual differences 107
2.6.3 The complexity of heterogeneous data 107
2.7 Evolving the Models 109
2.7.1 Linear Models 109
2.7.2 Generalized Linear Models 110
2.7.3 Generalized Linear Mixed-Effects Models 113
2.7.4 Random effects: intercepts and slopes 116
2.7.5 Selecting the best model 117
2.7.6 Disadvantages of the GLMM 117
2.8 Non-linear Mixed-Effects Models 118

3 Lateral inhibition in Migraine-with-Aura 121
3.1 Introduction 121
3.1.1 Altered Lateral Interactions in Migraine 123
3.1.2 Masking in Migraine 124
3.1.3 Global Processing in Migraine 124
3.1.4 Technique for assessing Lateral Interactions 125
3.2 Materials and Methods 126
3.2.1 Participants 126
3.2.2 Apparatus 128
3.2.3 Stimuli 128
3.2.4 Procedure 129
3.3 Results 130
3.3.1 Analysis 130
3.3.2 Between groups comparison of contrast 131
3.3.3 Between groups comparison of flanker facilitation 132
3.4 Discussion 133
3.4.1 Contrast Sensitivity 134
3.4.2 Inhibition and Lateral Interactions 135
3.4.3 Global Processes and External Noise 138
3.5 Conclusions 141

4 Spatial frequency tuning of perceptual learning and transfer in global motion 143
4.1 Introduction 143
4.1.1 The Reverse Hierarchy Theory 144
4.1.2 Experiment 1 147
4.1.3 Experiment 2 148
4.2 Predictions 150
4.2.1 Cortically localised global reweighting 150
4.2.2 Backward generalisation 150
4.3 Methods and Materials: 152
4.3.1 Participants 152
4.3.2 Experiment 1: 152
4.3.3 Experiment 2: 154
4.3.4 Procedure 156

4.4 Results 157
  4.4.1 Statistical Methods 157
  4.4.2 Feedback and Perceptual Learning: Experiment 1 159
  4.4.3 Feedback and perceptual learning: Experiment 2 161
  4.4.4 Transfer to global motion: Participant level variability 167
  4.4.5 Transfer to global motion: Summary 168
  4.4.6 Pre- and Post-Test Results for Contrast Sensitivity 168
4.5 Discussion 172
  4.5.1 Feedback and Perceptual Learning: Experiment 1 172
  4.5.2 Learning and Transfer of Global Motion: Experiment 2 176
  4.5.3 Summary of Training 176
  4.5.4 Evaluating the transfer from global motion 178
  4.5.5 Comparisons to Levi et al. 180
  4.5.6 Mechanisms of perceptual learning 182
  4.5.7 Models of perceptual learning 184
  4.5.8 Afterthoughts, Speculation and Unresolved Points 185
4.6 Conclusions 188
4.7 Acknowledgments 189

5 Local and global learning with and without feedback **191**
5.1 Introduction 191
  5.1.1 Supervised and unsupervised learning 193
  5.1.2 Easy and difficult trials 196
  5.1.3 Levels of processing 197
  5.1.4 Demands of different tasks and processing mechanisms 199
  5.1.5 Methods of presentation 203
  5.1.6 Aims and predictions 204
5.2 General Methods 206
  5.2.1 Participants 206
  5.2.2 Apparatus 206
  5.2.3 General Stimuli and Procedural information 206
5.3 Results 211
  5.3.1 General Analysis information 211
  5.3.2 Experiment 1 - Local and Global Form 214

5.3.3 Experiment 2 - local and global motion 220
5.3.4 Experiment 3 - Global motion using MOCS 226
5.4 Discussion 233
5.4.1 Predictions and Findings 237

6 Depth perception in anti-correlated random dot stereograms 243
6.1 Introduction 243
6.2 Psychophysical Experiment 255
6.2.1 Methods 255
6.2.2 Participants 255
6.2.3 Apparatus 256
6.2.4 Stimuli 256
6.2.5 Circular 257
6.2.6 Horizontal Edge 257
6.2.7 Vertical Edge 257
6.2.8 Procedure 258
6.2.9 Results 259
6.2.10 Depth Perception 259
6.2.11 Circular Stimuli 260
6.2.12 Horizontal Stimuli 261
6.2.13 Vertical Stimuli 261
6.2.14 Confidence 264
6.2.15 Relationship between Confidence and Performance 266
6.2.16 Discussion 268
6.3 Modelled responses to CRDS and ACRDS 271
6.3.1 Methods 272
6.3.2 First- and Second-Order Mechanisms 272
6.3.3 Stimuli 274
6.3.4 Procedure 274
6.3.5 Results 275
6.3.6 Discussion 278
6.4 General Discussion 280
6.4.1 Conclusion 287

7 Summary and Conclusion **289**
7.1 Summary of Aims 289
7.2 Summary of Empirical Results 294
7.3 General Discussion 298
7.4 Conclusion 302

References **305**

List of figures

1.1 Random-dot stereogram 37
1.2 Cross section of the retina 39
1.3 The visual hierarchy 42
1.4 V1 receptive cells 44
1.5 Local versus global processing 47
1.6 Local versus global task 51
1.7 Absolute Disparity 56
1.8 Relative Disparity 58

2.1 Sample psychophysical stimuli 88
2.2 The normal and cumulative distributions 91
2.3 The sigmoid shaped curve 91
2.4 Identifying the threshold from a psychometric function 92
2.5 Unreliable thresholds from psychometric function 94
2.6 The parameters of a psychometric function 99
2.7 Probability decision tree 102
2.8 Illustrative hierarchy of a data-set for a single participant 114

3.1 Experimental stimuli used to test lateral inhibition 129
3.2 Experimental Procedure 130
3.3 Contrast detection thresholds 133

4.1 Experimental Predictions 149
4.2 Global motion experimental stimuli 153
4.3 Illustrative psychometric functions 158
4.4 Feedback versus No-Feedback 160
4.5 Five Days Training Results 163
4.6 Pre and Post-Test Global Motion 166
4.7 Change scores (pre versus post) 169
4.8 Pre and Post-Test Contrast Sensitivity 170

5.1 Local and Global Processing 198
5.2 Eureka Effect 200
5.3 Local and Global Form stimuli 209
5.4 Local and Global Motion stimuli 210
5.5 Form Results 218
5.6 Local and global motion Results 224
5.7 Global Motion Results MOCS vs Quest 232

6.1 Binocular energy response 245
6.2 Correlated RDS and Anti-correlated RDS 248
6.3 Inverted response to ACRDS 250
6.4 A second-order binocular energy model 252
6.5 Da Vinci stereopsis 254
6.6 Stimuli and Procedure 258
6.7 Population level responses for CRDS and ACRDS 262
6.8 Response traffic light 264
6.9 Mean confidence for CRDS and ACRDS 265
6.10 Heatmap: Relationship between consistency and confidence 268
6.11 First- and second-order responses 277
6.12 Heatmaps for simulations 279

7.1 Schematic of global motion stimuli 300

List of tables

2.1 Alternative choices and probability of a correct response 90

4.1 Goodness of fit tests for each model. A mixed-effects model with the intercept as the only free parameter was compared with the model including the full psychometric function. -2 x the difference in log likelihood between the full and null model was calculated, and tested for significance against a chi-squared distribution, with degrees of freedom equal to the difference in the number of parameters for the two models. 160

4.2 Goodness of fit tests for each model. A mixed-effects model with the intercept as the only free parameter was compared with the model including the full psychometric function. -2 x the difference in log likelihood between the full and null model was calculated, and tested for significance against a chi-squared distribution, with degrees of freedom equal to the difference in the number of parameters for the two models. 164

4.3 Motion Coherence: Goodness of fit tests for each model. A mixed-effects model with the intercept as the only free parameter was compared with the model including the full psychometric function. -2 x the difference in log likelihood between the full and null model was calculated, and tested for significance against a chi-squared distribution, with degrees of freedom equal to the difference in the number of parameters for the two models. 167

4.4 Contrast Sensitivity : Goodness of fit tests for each model. A mixed-effects model with the intercept as the only free parameter was compared with the model including the full psychometric function. -2 x the difference in log likelihood between the full and null model was calculated, and tested for significance against a chi-squared distribution, with degrees of freedom equal to the difference in the number of parameters for the two models. 172

5.1 Local form training results across the three days for the measured thresholds at 65% and 85% accuracy, where the No-Feedback group are the reference group 215
5.2 Global form training results across the three days for the measured thresholds at 65% and 85% accuracy, where the No-Feedback group are the reference group 215
5.3 Local Form across session comparison of proportion of correct responses across stimulus intensity (angle of tilt) between the Feedback groups. The No-Feedback group are the reference group 217
5.4 Global Form: Comparison of proportion of correct responses across stimulus intensity (Coherence), by session within group, and between the Feedback groups, where the No-Feedback group are the reference group 219
5.5 Local motion training results across the three days for the measured thresholds at 65% and 85% accuracy, where the No-Feedback group are the reference group 221
5.6 Global motion training results across the three days for the measured thresholds at 65% and 85% accuracy, where the No-Feedback group are the reference group 222
5.7 Local motion Test: Comparison of proportion of correct responses across stimulus intensity (Degree of Tilt), by session within group, and between the Feedback groups, where the No-Feedback group are the reference group . . 223
5.8 Global motion test: Comparison of proportion of correct responses across stimulus intensity (Degree of Tilt), by session within group, and between the Feedback groups, where the No-Feedback group are the reference group . . 225
5.9 Global Motion Training results (MOCS). Comparison of proportion of correct responses across stimulus intensity (Coherence), by session within group, and between the Feedback groups, where the No-Feedback group are the reference group 228
5.10 Global Motion test (MOCS). Comparison of proportion of correct responses across stimulus intensity (Coherence), by session within group, and between the Feedback groups, where the No-Feedback group are the reference group 230
5.11 MOCS vs QUEST - test comparison 231

6.1 Fixed effects for near responses for CRDS and ACRDS 263

6.2 Fixed effects for confidence in CRDS and ACRDS judgements 266

Chapter 1

Introduction

1.1 An introduction to visual perception

Vision has been essential for the evolutionary survival of the human race, and is the sense most often relied on for everyday tasks; from navigating around a busy environment, recognising and interacting with objects, to sourcing food and avoiding threats (Tong, 2018). Having sight, and the ability to perceive the visual world seems so effortless that it can easily be taken for granted. The physiological processing of visual sensory information is well understood, starting at the retina of the eye as an electro-chemical chain reaction to light, where light-information is transduced into neural messages (Tong, 2018). These signals progress systematically through a sequence of specialised hierarchical brain areas that relay information about colour, motion, shape and depth. However, the power in visual perception lies hidden in the dense and complex information that our brains integrate in order to rapidly interpret and decipher the constantly changing visual environment. Along with the sensory information, processes such as attention, contextual and situational information, and prior experience, all contribute to this perceptual experience. To that extent, visual perception can be thought of as a subjective psychological experience the perceiver has when observing a scene.

1.1.1 The historical study of visual perception

The study of visual perception grew from thinkers of philosophy of mind. Early Greek philosophers approached perception through reflection, introspection and speculation. Aristotle described his experiences in vision, after administering a blow to his eye, and described in detail the aftereffects from staring directly at the sun (Wade & Swanston, 2013).

However, the blueprint for the study of visual perception dates back to the 17th century, when Rene Descartes distinguished between the psychological experience of perception and the physical properties of light (Bruce et al., 2003; Wade & Swanston, 2013). Descartes and other philosophical thinkers questioned the *'mind-body'* problem; that is, how are the physical body and “mental” related? (Miłkowski, 2018, p.74). Descartes compared the movements of people and those made by automated figures and from this, described learning mechanistically. He proposed that during an action, “associative memory traces" were imprinted on the brain (Stinson & Sullivan, 2018, p.376). This mechanistic view brought a new understanding to the study of perception, however he was not able to directly link how the mind and body interacted. Locke, at the end of the 17th century proposed that all knowledge was derived from one’s sensory experience; learned through observation and association (Zednik, 2018). Immanuel Kant disagreed with Locke arguing that perception was not a passive process, and suggested that people are born with an ability to perceptually organise *in space and time*. Kant's ideas went on to heavily influence Gestalt psychology (Wade & Swanston, 2013).

Discovering that the eye functioned in a similar way to a camera in the 18th century influenced the understanding of visual perception. Much like the time of exposure, it was thought that, over time, more information was added to the retinal image. However, it was thought that perception could not be measured, and very little in the way of experimental evidence was collected (Wade & Swanston, 2013). This changed when Gustav Theodore Fechner who, after lying in bed on a chilly October morning in 1850 pondered the complexities of sensation. He devised mathematical procedures to measure the transition of an event outside the body to the sensation that occurs inside the mind (Gescheider, 1997; Laming, 2011; Leek, 2001; Treutwein, 1995). Building on Weber's Law, which states that the just noticeable difference in sensation is proportional to the intensity of the stimulus, Fechner recognised there was a link between human behaviour and neurophysiology. Fechner coined the term *psychophysics* to describe the methodical approach to measuring and quantifying the relationship between sensory information and subjective perception (Gescheider, 1997; Kingdom & Prins, 2010, 2016). Fechner proposed that a sensory experience is proportional to the logarithm of stimulus intensity. Fechner's log function did not hold true in its specific details, with Stevens (1957) arguing that a power-law, rather than a logarithmic relationship, better accounts for the relationship between sensory experience and stimulus intensity. Nevertheless, the measurement techniques are still used in perceptual psychology today

(Lawless, 2013). By the 19th century psychology and sensory perception were examined experimentally, with new instruments through introspection rather than by observation.

1.1.2 Visual perception and the brain

In the mid 19th century Pavlov discovered his dogs would salivate, not only in the presence of food, but also as a result of a stimulus that regularly preceded food (Stinson & Sullivan, 2018). Pavlov proposed a causal connection between the external stimulus and physiological response where impulses in different areas in the brain are linked by the formation of novel connections (Stinson & Sullivan, 2018). While Pavlov conducted many behavioural experiments, he was not able to explain how and where the changes in neural connections occurred. Only when Camillo Golgi (1873) used silver nitrate stains on unmyelineated nerve cells, did an explanation on how nerve cells form begin to emerge (Stinson & Sullivan, 2018; Tong, 2018). From this, Santiago Ramón y Cajal discovered that neurons grow outwards and grow additional dendrites and axon collateral(s) over time. Cajal also identified that the nerve impulses travelled in one direction only, from the dendrites to the axon terminals. Furthermore, that the connections between nerve cells were separated by a hypothetical "intercellular barrier" Bennett (1999 cited by Stinson and Sullivan, 2018) that was so small that it was undetectable under a microscope. Along with Galvani's recently established finding that neural cells transmit electrical currents, speculation began to form that learning may involve increased branching of nerve fibres. Cajal suggested that to understand perception we need to understand the pathways of the neural circuitry in the vision processing areas (Stinson & Sullivan, 2018). The next piece of the puzzle was Sherrington's concept of the synapse (Stinson & Sullivan, 2018). He suggested that the barrier between each nerve cell acts like a valve to ensure transmission between nerve fibres only occurred in one direction and secondly, that there was a change in the characteristics of the impulse as it moved across nerve cells. In the early 20th century Hebb hypothesised that a metabolic change in growth occurred in cells when neural transmission (or firing) occurred in cells that were close in proximity, and provided an insight on methodologies to test for plasticity. By the 1960s, scientists had fully modelled the structure of a neuron, and with the advancement in techniques of cellular neuro-physiology Kandel and Spencer (1968) considered approaches to studying learning and neural plasticity. The detailed review provided several approaches to measuring plasticity behaviourally. In 2000 Eric Kandel was co-awarded the Nobel Prize for Physiology for his work in the marine mollusc (Aplysia californica) (Aggio & Derby, 2010), for his approach to how neurotransmitters function in the brain (Aggio & Derby, 2010; Kandel, 2004). The nervous system of the Aplysia has a small number (20 000) of very large neural cells, making it more accessible than the human brain's million-million neurons (Kandel, 2004). By administering tactile stimuli to the Aplysia, over a period of time, three types of neurons were identified; the sensory, motor and interneurons. Furthermore, Kandel (2004) showed that the Aplysia 'remembered' previous exposure to the stimulus, there-by distinguishing between the mechanisms that lead to long or short-term memory (Kandel, 2004). Experience was found to change the strengths and effectiveness of the chemical connections between the nerve cells, and thus synaptic plasticity emerged as the mechanism for learning in the nervous system (Milner et al., 1998).

1.1.3 The modern approach to understanding visual perception

More recently, visual perception is understood through exploring three important characteristics of visual processing; i) functional specialisation, ii) hierarchical organisation and finally iii) the neural connections between each functional area. As a result of physiological evidence, largely supplied by the Hubel and Wiesel cat studies (1972; 1977; 1978), it was generally accepted that the visual brain was organised in a hierarchical, successive manner. However, Zeki (2003) note that Hubel and Wiesel's studies were undertaken primarily using cats and kittens, that have a visual system that is different from humans, for example by lacking colour vision. However, later studies of the primate visual cortex revealed the presence of functionally specialised visual areas (Livingstone & Hubel, 1988; Zeki, 1978). The characteristics of functional areas is inferred, based on the responses of neurons at that level (Lennie, 1998). Each cell has a preferential attribute (or group of attributes) for which it responds to. The cells that are involved in processing and perceiving an attribute are said to be located in functional brain areas (Lennie, 1998; Zeki, 2003). There is significant evidence for the specialisation of neural processing of visual information. For example with, complex motion in V5 (Britten et al., 1993; Maunsell et al., 1990; Newsome et al., 1989; Rudolph & Pasternak, 1999), object and colour recognition V4 (Grill-Spector et al., 2001; Roe et al., 2012; Zeki, 1983), simple motion in V3 (Felleman & Van Essen, 1991) and disparity tuning for binocular vision in V2 (Assee & Qian, 2007; Bredfeldt & Cumming, 2006; Bredfeldt et al., 2009; Thomas et al., 2002).

The function and size of cells differ based on their specialisation and location. The receptive fields of cells at lower cortical levels (such as V1) are very small and only respond to a limited area of the visual field, and to specific attributes (or features) (Furlan & Smith, 2016; Lamme, 2003; Simoncelli & Heeger, 1998). However, at higher levels, where processing is more complex, cells respond to groups of attributes. Receptive fields in higher areas (such as V5) need to integrate responses from a wider area of the visual field, and pool the responses from several lower level cells (Amano et al., 2009; Bex & Dakin, 2002; Burr & Thompson, 2011; Gilbert et al., 2001; Nishida, 2011). Higher up the visual hierarchy, areas can become more specialised and have very specific responses to complex stimuli; for example the fusiform face areas, are known to only respond strongly for faces (Blake & Wilson, 2011).

Structural and functional areas are interconnected by complex neural circuitry (Callaway, 2004). There are numerous visual processing pathways that distribute information throughout the levels of the processing hierarchy (Felleman & Van Essen, 1991). However, it is important to note that information does not flow in a simple, feedforward manner. In addition to the feedforward connections there are also feedback and lateral connections (Callaway, 2004). It has been suggested that while the feedforward connections provide a rapid automatic response, the reiterative input from feedback connections provide the perceptual detail (Hochstein & Ahissar, 2002; Jehee et al., 2007; Lamme, 2003; Pascual-Leone & Walsh, 2001). Recent evidence has supported this proposal where the feedback connections from V5 to V1 were found to be important in the perception of motion (Romei et al., 2016a).

Furthermore, it has been proposed that the feedback connections may play an important role in learning (Ahissar & Hochstein, 2004; Hochstein & Ahissar, 2002).

Each of these characteristics are discussed in more detail, within the introduction and the thesis.

1.1.4 Neural plasticity and perception

Neural plasticity describes the brain's ability to change, following an experience. The result is a permanent change as a result of new or modified synaptic connections (Milner et al., 1998). Plasticity is a fundamental requirement for learning and memory, functional development and for recovery after a brain injury or trauma (Kandel, 2004; Lunghi & Sale, 2015; Milner et al., 1998). Neural plasticity is most robust during childhood, during a period shortly after birth, known as the critical period, and declines in adulthood (Gilbert et al., 2001; Hensch, 2005).

Research into the mammalian visual system in the 1950s suggested that the neural pathways responsible for visual processing are hard-wired in adulthood (Karni & Sagi, 1991). Visual sensory development was thought to be restricted to the critical period (Gilbert et al., 2001), the stage for which experience (i.e. sensory experience) provides the information that is essential for the development of the neural systems that process the experience (Hensch, 2005). The critical period for the normal development of visual function has been investigated using visual deprivation in kittens (Hubel & Wiesel, 1970; Wiesel & Hubel, 1963; Wiesel et al., 1963), rats (Nealey & Riley, 1963) and non-human primates (Reisen, 1947).

Hubel and Wiesel initiated some of the earliest physiological studies into plasticity of the visual cortex. Their studies of the visual system (Hubel & Wiesel, 1962) launched a new field of research and established the general principles that underpin neural development in visual perception (Daw, 2009). Single-cell responses in kittens deprived of light in one eye (through covering the eye or stitching the eye-lid) for the first 1-4 weeks of life were recorded (Wiesel & Hubel, 1963). The Lateral Geniculate Nucleus (LGN) relays sensory information from the retina to the primary visual cortex, with alternating layers of monocular cells responding to the two eyes. There was a significant reduction in the number of cells in the LGN corresponding to the occluded eye, however the cells in the primary visual cortex were found to be mostly normal (Wiesel & Hubel, 1963; Wiesel et al., 1963).

Once the sutures or covers over the eye were removed, kittens deprived for 2.5 months displayed ocular dominance for the healthy eye only, with no input from the deprived eye. However, atrophy was significantly reduced if the kitten had the eye sutured at the age of 9 weeks, for periods as long as four months. In this case responses to input from the healthy eye dominated, however the deprived eye was also found to influence responses more than in those kittens deprived from birth. Finally, when the same procedure was conducted using an adult cat there was no atrophy or deterioration to the LGN after 3 months. The researchers concluded that monocular deprivation produced permanent defects in the visual system, since the cells no longer responded to sensory information stemming from the deprived eye. Furthermore, since the majority of cortical cells are binocular, they proposed that the functioning of the primary visual cortex was driven by the responses from the healthy eye in monocularly-deprived cats, and since these deficits were not visible when deprivation occurred in adulthood, the defects are evidence for a critical growth period of development shortly after birth (Wiesel & Hubel, 1963; Wiesel et al., 1963). Depriving kittens of light in one eye for the first three months of life was found to permanently compromise binocular vision (Wiesel & Hubel, 1965). Eye closure had little effect when conducted within the first 3 weeks of life, while after the third week, any sensory deprivation within the first three months was found to permanently damage development; kittens deprived for as little as 30 days never developed binocular connections (Hubel & Wiesel, 1970).

The development of depth perception under conditions of deprivation has also been investigated behaviourally in rats. Rats were kept in the dark for varying lengths of time to establish how much deprivation was required for them to lose the ability to perceive depth (Walk et al., 1965). Dark-reared rats exposed to light within 140 days of deprivation were able to perceive depth, however those deprived for longer were not. Similarly, the critical period in non-human primates has been found to occur within the first sixteen months (Reisen, 1947), with similar outcomes to those found in kittens. From their work on deprivation, Wiesel and Hubel (1965) concluded that where sensory experience during the critical period was abnormal, damage was permanent. Thus it was predicted that the human brain was shaped during early childhood, and could not be modified in adulthood (Sagi, 2011).

1.2 Perception and Learning

The evidence for the critical period of the development of neural circuitry, implied that the visual system was fixed in adulthood (Gilbert et al., 2001; Karni & Sagi, 1991). More recently, it has been established that, rather than a single fixed duration for sensory development, there are multiple periods of development for different visual functions (see Lewis and Maurer (2005) for a review of the multiple sensitive developmental periods for visual development in children). However, the critical period for children generally occurs within the first few (1-2) years of life (Lewis & Maurer, 2005). However, it has been argued that the presence of a critical period does not preclude the potential for learning in an older brain (Hensch, 2004). This is evident in perceptual learning, which naturally contributes to our ability to adapt to and navigate within an ever-changing environment.

1.2.1 Defining perceptual learning

Perceptual learning has been a topic of psychological research for over 150 years and is one of the earliest topics in the study of perception (Gold & Watanabe, 2010). Perception was theorised to occur as a gradual learning process where the brain combines accumulating memory traces from each sensory experience (Stinson & Sullivan, 2018). Thus, perception was considered to be an additive process whereby sensory experiences would accumulate over time and gradually enrich the ability to perceive (Gibson & Gibson, 1955).

However, Gibson and Gibson (1955) questioned the definition of perceptual learning. Is human perception a learned process? Where for example, humans learn to see depth, or form, and take meaning from these observations. Or, is human perception the driving force behind learning? Where understanding the world largely depends on perception (Gibson & Gibson, 1955)? Posing two questions, “do we learn to perceive" or “can we learn by perceiving" (Gibson & Gibson, 1955, p.32) undertook an empirical study using a simple discrimination experiment. Adult and child participants were shown an image of a scribble after which they were required to identify it, from a batch of approximately 20 cards. No feedback was provided, but after each run they were allowed to look at the original card. By the third run, performance was perfect for adults, while the children's performance was more variable, older children (8-11 years) achieved perfect scores in around 5 attempts. In their experiment learning increased the likelihood of a correct response, but while repetition was necessary it was argued that there was no evidence for memory (Gibson & Gibson, 1955).

Based on a traditional view of perception by enrichment, Gibson and Gibson (1955) argued that perception is a creative process, building progressively from images stored in memories, that accumulates by association to a sensory cue (Adolph & Kretch, 2015; Gibson & Gibson, 1955). However, the results from their study suggested perceptual learning, was not about gradually adding memories over time, in an imaginatory process to enhance sensory experience. Rather, more exposure increases the ability to discriminate (Gibson & Gibson, 1955). Perceptual learning as a process improves the likelihood of discrimination, differentiation and extraction of information that was always there, but not previously detected (Adolph & Kretch, 2015).

Gibson (1963) defined *perceptual learning* as a *fairly* permanent improvement in perception as a result of experience or practice. This way of thinking motivated a new interest in the topic of perceptual learning.

1.2.2 Differentiating perceptual learning from other types of learning

Learning can be divided into two broad categories, associative learning or non-associative learning (Byrne & Hawkins, 2015; Hawkins & Byrne, 2015a). Associative learning describes the pattern of learning obtained through behavioural conditioning, such as seen with Pavlov's dog. Learning occurs because the observer learns that there is a relationship between two independent events that occur in tandem (Hawkins & Byrne, 2015b). Non-associative learning includes habituation and sensitisation (or sensory adaptation) (Byrne & Hawkins, 2015). Habituation is the phenomenon that describes a diminishing response to a repeatedly presented stimulus. In contrast, sensitisation occurs in an opposite manner, the repeated presentation (or exposure) of a stimulus results in an increased sensitivity to that stimulus (Byrne & Hawkins, 2015).

Associative learning is best described by a single characteristic, that the neural representation of multiple elements can be linked so that excitation of one also excites its associate as a conditioned reflex (Hall, 1991). However, the mere co-occurrence of two events was argued to be inadequate to explain the complexities of learning by conditioning (Wilson, 2012). The Rescorla-Wagner model emerged in the early 1970s to attempt to address these questions (Rescorla, 2008).

As a new addition to associative modelling the Rescorla-Wagner model included values to represent i) the salience of the stimulus and ii) the speed at which learning occurs (Hall, 1991; Rescorla, 2008; Rescorla & Wagner, 1972; Wilson, 2012). For example, the model recognises that when the perceiver is surprised by an experience, then learning is more likely.

Learning is less likely when the response is predicted. Furthermore, the expectation on any trial will depend on what has previously been learned (Hall, 1991). If you have no experience with a conditioned stimulus you predict and expect nothing. However, if the unconditioned stimulus occurs and you are surprised then you have learned something (Wilson, 2012). Thus, the model accommodates a stimulus that may contain different rates of associative strength (Rescorla & Wagner, 1972). The speed at which learning occurs, and the amount of learning is dependent on the amount of surprise (Wilson, 2012). Surprise is calculated as the difference between what was expected and what happened (Hall, 1991; Rescorla, 2008; Rescorla & Wagner, 1972; Wilson, 2012). The Rescorla-Wagner model (Rescorla & Wagner, 1972) is considered the most influential model of associative learning (Hall, 1991).

There are numerous models of associative learning and each model attempts to formalise in greater depth, the conditions for which associations are formed (Hall, 1991). The key characteristic in all the models, is that there is simultaneous activation of the neural representations between a pair of events, and this co-activation strengthens their link. While there has been some disagreement as to the details (see (Hall, 1991)), associative theorists agree that strengthened associations are a result of pairing the level of activity of a conditioned stimulus and the level of activity in the unconditioned stimulus (Hall, 2003). The magnitude of the co-activation will depend on the strength of the stimulation, such as the salience of the conditioned stimulus (Hall, 2003).

Several researchers (Hall, 1991; Mclaren & Mackintosh, 2000) have proposed associative learning models that attempt to account for perceptual learning. While not all perceptual improvements are explained by perceptual learning, the differentiating feature of the latter is that perceptual learning occurs independently of associations between other cognitive factors (Gold & Watanabe, 2010; Hall, 2003). Perceptual learning can be influenced by mere exposure to a stimulus, whereby repeated experience changes its neural representation and enhances discriminability, in a way that is unexplained by models of associative learning (Hall, 2003). Associative learning establishes the neural weights between a sensory input and a stimulus as a function of neural tuning (Law & Gold, 2009). In contrast, perceptual learning shapes the weights so that the neurons responding to those stimuli are weighted more strongly (Law & Gold, 2009).

However, if categorising perceptual learning as non-associative learning, it is important to note that it is also distinct from other forms of non-associative learning such as sensory adaptation and habituation (Gold & Watanabe, 2010). Habituation and sensory adaptation (or sensitisation) are short term learning mechanisms that result in a temporary change in the response to a repeatedly presented stimulus (Byrne & Hawkins, 2015; Gold & Watanabe, 2010). In contrast, perceptual learning has been evidenced to be permanent or relatively stable (Fahle, 2002; Gold & Watanabe, 2010). For more information, Hall (2003, p.50) provides a thorough discussion on adaptation and habituation in relation to associative learning.

Perceptual learning is long-lasting, occurs outside of conscious awareness, is robust in both children and adults, and does not require behavioural conditioning. As such it is argued that perceptual learning reflects an intrinsic attribute of the perceptual system (Fahle, 2005; Gold & Watanabe, 2010; Sasaki et al., 2010).

1.2.3 Measuring Perceptual Learning

Perceptual learning is often studied in a similar manner as perception, where participants repeatedly (sometimes across days) perform the same task(s). Performance on tasks that require detection, discrimination or categorisation is monitored, measured and evaluated. Ultimately perceptual learning demonstrates that with effective and appropriate training people improve on most perceptual tasks. For example perceptual training has been used by researchers to train novice observers to decipher visual cues for tasks that would be otherwise uninterpretable, such as chicken sexing (Biederman & Shiffrar, 1987), detecting cancer in mammograms (Kundel et al., 2007) and locating abnormal features in an x-ray (Sowden et al., 2000). In each of these examples, a task that was impossible for a novice, can be successfully performed after training. Perceptual learning has since been found to occur in every sensory modality (Dosher & Lu, 2017). The responsiveness of the adult visual system to learning suggests that the perceptual system is malleable (Lu et al., 2011). This is in stark contrast to the early understanding that sensory development is restricted to a critical period early in adolescence, and that the visual system in the adult brain is hard-wired (Hubel & Wiesel, 1970; Sagi, 2011; Seitz & Watanabe, 2005).

1.2.4 Perceptual learning in early vision research

Early in perceptual research, observers were often provided with practice trials in order to familiarise them with the experimental task (Sagi, 2011), and some tasks were only possible after a period of training. However, only a few publications directly reported on the effects of training. One of the earliest reports of perceptual learning (Sagi, 2011) is found in the field of binocular vision. Binocular neurons receive slightly different images from each eye, and use these to create the perception of the three dimensional world. The perception of binocular depth can be simulated in a laboratory by using random-dot stereograms (Julesz, 1971). Random-dot stereograms, (see Figure 1.1) are comprised of random dots presented to each eye, with a shift in the positions of individual dots to one eye, that simulates the perception of stereoscopic depth. However, these tasks are notoriously difficult for inexperienced observers to fuse (Julesz, 1971; Sagi, 2011). Successful fusion was found only to occur after practice. In a similar task, where the stimuli were comprised of fine orientated line segments, post training improvement was found to be selective to the orientation of the trained stimulus (Ramachandran & Braddick, 1973). The specific improvement in orientation suggested that training may have stimulated 'orientation analysers' (Ramachandran & Braddick, 1973, p.375). This is proposed to be one of the earliest links made between neural mechanisms in the visual cortex and specific perceptual learning stimuli (Sagi, 2011). Training induced improvements were shown to occur in very early studies for tasks such as reducing detection time for visual stimuli (Ramachandran & Braddick, 1973) or increasing sensitivity when discriminating small differences in the position of two lines (Vernier acuity) (Mckee & Westheimer, 1978). Fiorentini and Berardi (1980) evaluated if observers trained to discriminate oriented gratings retained their improved ability over time. Secondly, they were interested in identifying if improvements in one orientation and spatial frequency improved performance on untrained orientations and spatial frequencies. After a short training period, the ability to discriminate the complex gratings (with varying luminance profiles) improved.

Improvement was retained from one day to the next and importantly changing the spatial frequency or re-orienting the gratings by 90° removed any training effects (Fiorentini & Berardi, 1980). With an increasing number of studies reporting improvements, researchers began to question if there was an element of plasticity in the nervous system (Fiorentini & Berardi, 1980). Perceptual learning became an established method to improve the ability to detect or discriminate previously unseen stimuli (Fahle, 2002). Perceptual changes had previously been attributed to high level visual cognition (Fahle, 2002). However, the reported improvements after training were routinely found to be highly specific to low-level features of trained stimuli (Fahle, 2005; Sagi, 2011). Low-level features are those features such as the position and orientation of a stimulus (Fahle, 2005; Fiorentini & Berardi, 1980), but can also include features such as the eye to which it has been presented (Karni & Sagi, 1991). On the basis of these low-level, highly specific improvements, early attempts to identify the neural substrate of the learning process focused on the early visual areas, where neurons are tuned to stimulus dimensions such as orientation, position and spatial frequency (Fiorentini & Berardi, 1980). This led to the presumption of a low-level neural locus of the perceptual learning mechanism (Gilbert et al., 2001).

Since then, many psychophysical studies have demonstrated that, with practice, thresholds for detection and discrimination of visual stimuli can be reduced (improved) (Seitz & Watanabe, 2005). Furthermore, perceptual learning has been observed to occur in many visual detection or discrimination tasks (Fahle, 2002), including contrast sensitivity (Polat, 2009; Sahraie et al., 2006; Sowden et al., 2002), orientation discrimination (McGovern et al., 2012), global motion (Huxlin et al., 2009; Levi et al., 2015; Sundareswaran & Vaina, 1995),

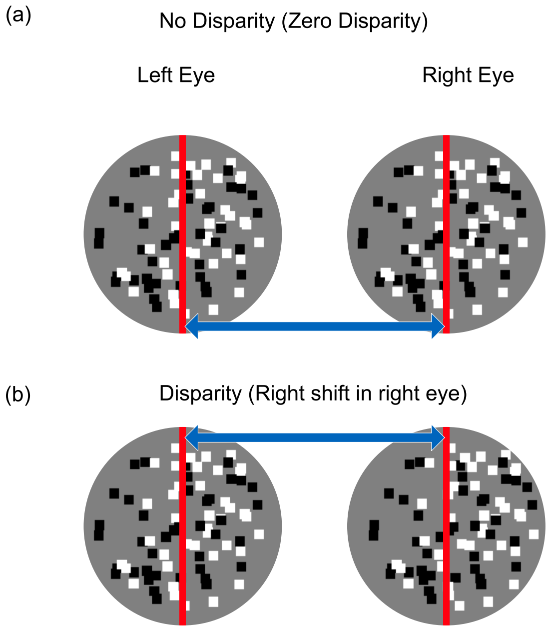


Figure. 1.1 Random-dot stereograms (RDS) are often used in laboratory settings to investigate binocular depth. The left image is presented to the left eye and the right to the right eye. The red line across the central vertical point is to illustrate the relative matching of the dots for each eye's image. The targets shown in (a) are at zero disparity and thus no depth would be perceived. For the targets shown in (b) there is a rightward shift in the image projected to the right eye providing the illusion of depth.

direction discrimination, (Ball & Sekuler, 1982, 1987; Huang et al., 2007; Seitz et al., 2006), discrimination of texture (Karni & Sagi, 1991, 1993; Schwartz et al., 2002) and hyperacuity (Fahle & Edelman, 1993; Levi et al., 1997; Saarinen & Levi, 1995; Weiss et al., 1993).

The responsiveness to training programmes suggests that even in adulthood, the perceptual system is able to learn (Lu et al., 2011). The benefits accrued from training have been found to be long-lasting, without the need for retraining (Maniglia et al., 2016; Petrov et al., 2005; Polat et al., 2004; Sagi, 2011). Perceptual training has been shown to produce improvements in perception across all perceptual modalities, including touch (Dinse et al., 2003), taste and smell (Chu et al., 2016; Wilson & Stevenson, 2003), hearing (Harrison et al., 2005) and vision (Karni & Sagi, 1991). These improvements are not limited to laboratory studies, and have real world benefits for individuals, such as improving radiographers' detection rates for breast cancer in mammograms (Kundel et al., 2007), robust improvements in visual perception in air force pilots with presbyopia (Sterkin et al., 2018), athletes (Deveau et al., 2014), medical professionals (Kellman, 2013) and people with dyslexia (Ahissar, 2007). Perceptual learning provides a valuable insight into how the human visual system functions (Seitz & Dinse, 2007), and while there are studies that have failed to report post training improvements in perception (Herzog et al., 2012; Herzog & Fahle, 1997; Lu & Dosher, 2004; Sagi, 2011; Seitz et al., 2006; Yu et al., 2004), these contribute to understanding how the separate systems may be involved in the learning process. Thus, an important theoretical question concerns the neural loci within the visual system of the learning mechanism (Asher et al., 2018b; Ball & Sekuler, 1987; Fiorentini & Berardi, 1980; Maniglia & Seitz, 2018; Owsley, 2011).

1.3 The functionally organised visual system

The visual system processes sensory information through a process of hierarchically organised structures starting at the retina. Ganglion cells (see Figure 1.2(a)) receive signals from the bipolar cells and provide the output signal to the magnocellular or parvocellular divisions of the Lateral Geniculate Nucleus (Livingstone & Hubel, 1988). Most retinal ganglion cells are approximately circular and display a centre-surround antagonism. That is, there are interactions between the regions of the receptive fields of these cells. For example (see Figure 1.2(b)), shows on (depicted by $+$)-centre and off (depicted by $-$)-centre receptive cells.

Regions that prefers light (+) or dark (−), are surrounded by off-surround and on-surround regions respectively. The on-centre and off-surround antagonistic responses provide key signalling for the detection of edges (Livingstone & Hubel, 1988).

1.3.1 Pre-cortical area : The Retina

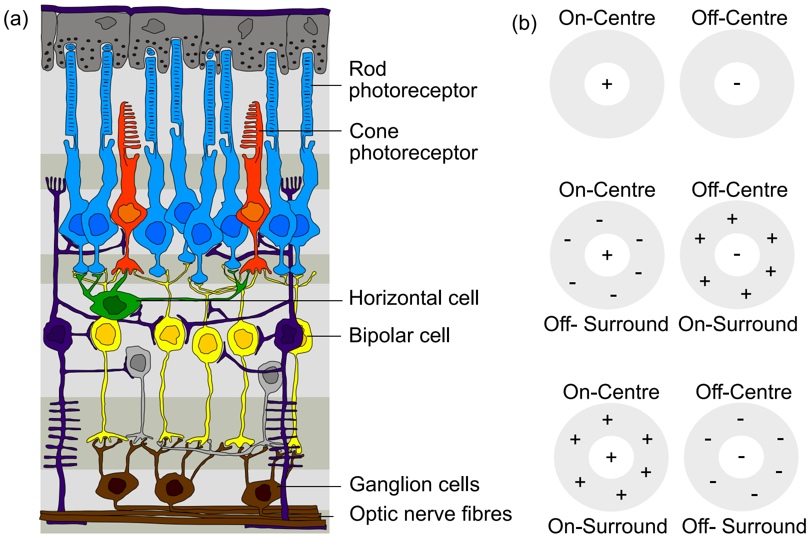


Figure. 1.2 (a) Diagram showing a cross section of the retina. Figure shows the Rod and Cone photoreceptors, the horizontal and bipolar cells and finally the ganglion cells and optic fibres. (Modified from Wikimedia Commons CC BY-SA 3.0) (b) On-centre and off-centre retinal ganglion cells. The centre-surround configuration of the ganglion cells transmits signals that include exposure to light as well as the varying firing rates of the centre and surround areas. Here, the plus and minus signs represent light stimulating the centre and surround. When light of the preferred polarity is presented only to the centre (top), or to both the centre and surround (centre), the neuron will respond. When light of the non-preferred polarity is presented to the surround (bottom) this will cancel the response from the centre, resulting in a weak response.

1.3.2 Sub-cortical area : Lateral Geniculate Nucleus

Information is relayed from the ganglion cells in the retina to the primary visual cortex via the LGN in the thalamus. Visually sensitive neurons in the LGN are organised in magnocellular, parvocellular and koniocellular pathways (Maunsell et al., 1990). The receptive cells in the LGN display the centre-surround structure of the retinal ganglion cells.

Magnocellular and parvocellular pathways

There is some evidence that specific groups of neurons are involved in perceptual processing, and there is accumulating evidence that there are areas of specialisation within the visual system (Parker, 2007). Some of the earliest evidence of this organised separation was found within the LGN where neurons are comprised of at least three channels of visual input (Maunsell et al., 1990). These cells are distinguished by anatomical and physiological differences and are referred to as the magnocellular, parvocellular and koniocellular pathways (Maunsell et al., 1990). The physiology of the magnocellular and parvocellular pathways has been studied in detail across primates and non-primates (see Xu et al. (2001) for review) and are thought to associate most strongly with the dorsal (where) and ventral (what) pathways (Goodale & Milner, 1992) respectively (Parker, 2007). The parvocellular neurons respond primarily to colour and fine detail: 90% of cells are sensitive to differences in wavelength and have a high spatial resolution (Livingstone & Hubel, 1988). Parvocellular cells are specialised for fine detail, textures and patterns, and are thought to provide the major input to the ventral system for object identification and border detection (Croner & Kaplan, 1995; Livingstone & Hubel, 1988; Maunsell et al., 1990; Merigan et al., 1991; Wiesel & Hubel, 1966).

Magnocellular neurons in the LGN have receptive fields that are larger than those of parvocellular neurons (Croner & Kaplan, 1995). Magnocellular cells have been found to respond more to moving and flickering stimuli and have high sensitivity to contrast, and changes in light onset (Croner & Kaplan, 1995; Kaplan & Shapley, 1982). Layer 4Cα is a sublayer in layer 4 of V1 (see Figure 1.3), and an area hypothesised to be the neural locus of contrast discrimination (Livingstone & Hubel, 1988; Sowden et al., 2002). 4Cα receives its primary input from the magnocellular layers of the LGN (Hubel & Wiesel, 1972). In addition to the independent motion pathway from LGN to V5 (Bridge et al., 2015), the magnocellular/dorsal pathway is also involved in the early processing of motion (Livingstone & Hubel, 1988). Single cell recordings from area V5 in the macaque found reduced responses after inactivation of the magnocellular channels in the LGN (Maunsell et al., 1990). The most important input to area V5 is argued to be the magnocellular projection via layer 4B of V1, where the receptive fields are selective for direction of motion and orientation but not for colour. Based on the physiology of the macaque it was originally proposed that binocular stereopsis was also processed exclusively in the magnocellular pathway/dorsal stream (Livingstone & Hubel, 1987). Livingstone and Hubel (1988) illustrated how the perception of depth is compromised when a coloured image is presented on a background that has been manipulated so as to only stimulate colour sensitive mechanisms (Lu et al., 1999). Since the magnocellular pathway does not respond to colour, this was taken as evidence that the magnocellular cells alone process stereo perception (Sincich & Horton, 2005). However, given that lesions applied to magnocellular layers in the macaque LGN were found not to disrupt stereo perception (Schiller et al., 1991), and physiological evidence for disparity tuned cells in areas outside of the magnocellular pathway, (see (Sincich & Horton, 2005)), it is now generally accepted that both pathways process stereo information (Parker, 2007).

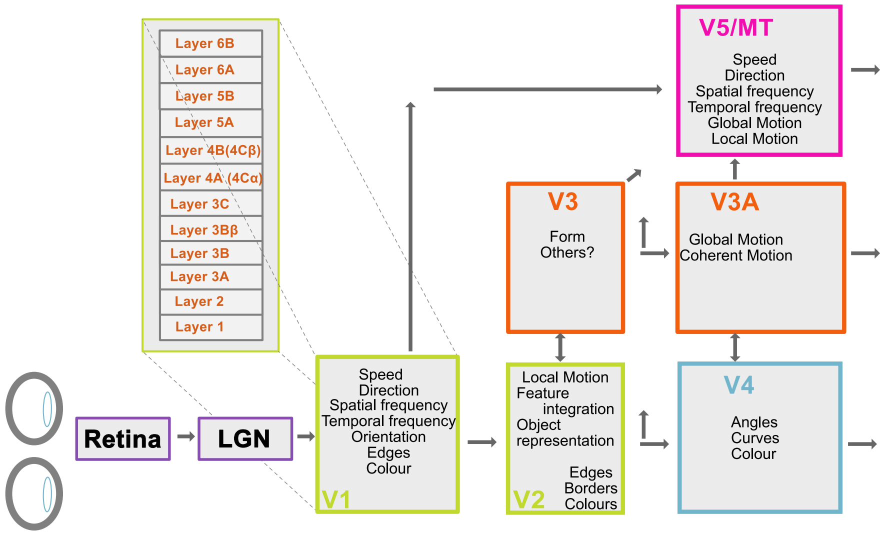


Figure. 1.3 The visual system is often represented as an interconnected, feedforward hierarchical network of functionally specialised areas. Typically the processing of sensory information begins at the retina, and proceeds to the LGN which is the relay system for visual information. The LGN segregates the sensory information for processing by more specialised areas. Area V1 is the largest of the visual areas, and it contains several specialised sub-layers. Receptive fields are small and specific for very basic features, such as orientation of edges and lines. V3, V4 and V5 are higher cortical areas and neurons in these areas are larger, show less specificity to individual features and ‘pool’ or average sensory information across a range of low level receptive fields.

1.3.3 The cortical hierarchy

The cortical areas of the visual system are typically represented as a hierarchy of interconnected areas (Lennie, 1998; Livingstone & Hubel, 1988) (see Figure 1.3). The LGN acts as the relay centre for the visual pathway. The visual system is broadly divided into multiple visual areas (Zeki, 1978). Each area is comprised of highly specialised receptive fields and this becomes more functionally specific higher in the hierarchy.

Receptive fields of a neuron

Receptive fields of a neuron represent a defined region in the visual field which elicits a response when stimulated (Sincich & Horton, 2005). The receptive fields at the lowest level areas, V1 and V2, represent simple features over a small area of the visual field and are then passed to the higher cortical areas such as V3, V4 and V5. The receptive fields of these areas are larger and pool the sensory information across several lower level receptive fields (Burr & Thompson, 2011; Gilbert et al., 2001). The strength of response of the neuron depends on how well the stimulus is matched to its receptive field. In early areas, receptive fields are tuned to simple properties such as orientation and scale. Cells in these areas are also tuned to properties such as direction of motion and colour. Receptive fields in higher areas tend to be larger, more specialised, more broadly tuned to dimensions such as scale, and respond to more complex features.

The primary visual cortex

The first cortical visual processing level is the primary visual cortex (V1). It is the largest of all visual areas of the cerebral cortex in the macaque (Felleman & Van Essen, 1991), and is comprised of six sub-layers (see Figure 1.3). V1 is retinotopically organised and incoming projections from the LGN are organised by the eye of origin. V1 extracts information of ocular origin, orientation tuning and visual space and organises these in columnar functional maps (Callaway & Nassi, 2009). The receptive fields in V1 are small and only respond to a very restricted area of the visual field (Furlan & Smith, 2016; Lamme, 2003; Simoncelli & Heeger, 1998). Simple cells (Figure 1.4(a)) sum the inputs from an aligned array of LGN receptive fields to build elongated receptive fields. The receptive fields resemble Gabor functions with odd-symmetric or even-symmetric arrangements of on- and off-regions (Figure 1.4(b)). These cells represent simple visual dimensions, such as the position, orientation and scale of local image features (Ahissar & Hochstein, 1997). The specificity of cells in primary visual cortex was identified using micro electrodes to record the activity of cells in cats (Hubel & Wiesel, 1959). The position and orientation of the stimulus was crucial in obtaining a strong reading from a specific neuron. Physiological evidence from a range of animals has shown that neurons in the primary visual cortex are tuned to orientation (DeValois et al., 1982; Hubel & Wiesel, 1962), spatial frequency (DeValois et al., 1982; Movshon et al., 1978) and direction (Livingstone & Hubel, 1988; Movshon & Newsome, 1996). Consequently, this has also been supported by electrophysiological findings, from the study of area V1 in monkeys (Gilbert et al., 2001) and functional magnetic resource imaging (fMRI) in humans (Seitz et al., 2005).

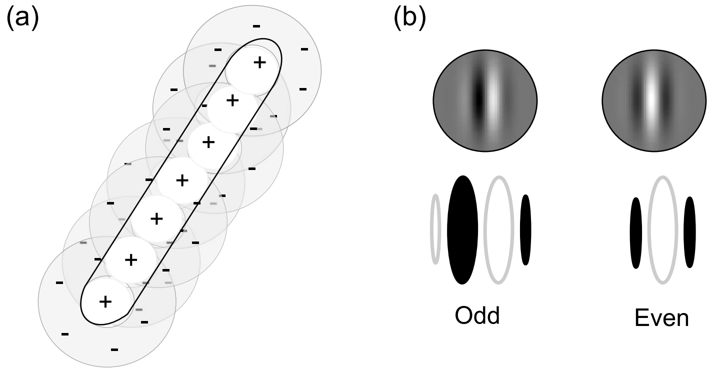


Figure. 1.4 (a) Simple cells in V1 sum the inputs from an array of LGN receptive fields that resemble an elongated bar with an on-centre and off-surround. (b) These receptive fields can be modelled using odd-symmetric and even-symmetric Gabor functions fields

Area V2 is the second largest visual area in the macaque, and receives two thirds of its input from V1 (Sincich et al., 2003). V2 plays an important role in the processing of orientation, colour, depth (Sincich & Horton, 2005) and motion (Hu et al., 2018). Recently, area V2 has also been shown to contribute to binocular depth perception (Hu et al., 2018).

Psychophysical experiments are also used to investigate the behavioural responses of receptive fields. Local visual tasks are defined as those that can be performed reliably based on the type of information which individual cells in early cortical areas such as V1 encode. These include the detection of stimuli and the discrimination of their orientation, position and direction of motion.

Extrastriate Visual Areas

The higher levels in the visual pathway, such as V3, V4 and V5, have neurons with much larger receptive fields. The tuning properties of these neurons are more complex, and their responses are less dependent on the location and retinal size of stimuli, the viewpoint of the observer and the prevailing lighting conditions (Felleman & Van Essen, 1987; Furlan & Smith, 2016; Hubel & Wiesel, 1965; Mikami et al., 1986; Movshon et al., 1978; Sillito et al., 2006; Zeki, 1974). The receptive fields of neurons in higher cortical areas integrate information to represent global stimulus properties (Amano et al., 2009; Bex & Dakin, 2002; Burr & Thompson, 2011; Gilbert et al., 2001; Nishida, 2011) across a collection of V1 receptive fields. At each step up the hierarchy the receptive cells get progressively more complex in their tuning. Higher cortical areas, beyond V4, begin to display category sensitive tuning. The fusiform face area (FFA) is a cortical area which is known to respond preferentially to faces (Blake & Wilson, 2011). Furthermore, object-specific cells have been identified in the lateral occipital cortex (Grill-Spector et al., 2001).

In psychophysical and physiological studies, global processes are investigated using stimuli or tasks that can only be resolved through integration and segregation of coherent or conflicting information (Burr & Thompson, 2011; Garcia et al., 2013; Nishida, 2011). A canonical example of this is the global motion stimuli introduced by Williams and Sekuler (1984). These stimuli contain dots moving in multiple directions, and the observer's task is to determine the direction of motion of the stimulus as a whole. Information in local regions of the image (and thus also the responses of V1 neurons) will reflect the direction of motion of individual dots, and the task can only reliably be completed by integrating information about the directions of all the dots across space. Cortical cells higher in the processing hierarchy are involved in the perception of global aspects of a percept and generalise across individual features such as spatial frequency, speed, motion and orientation. Receptive cells in the global processing levels pool sensory information from a selection of lower level receptive cells (see Figure 1.5).

1.3.4 Feed-forward and feed-back connections

The visual hierarchy is not organised on an exclusively feedforward (bottom-up) manner, and feedback connections are argued to be fundamental to efficient cortical organisation (Rockland & Knutson, 2000). V1 sends most of its feedforward output to V2, however there are also direct feedforward connections to V3 and V5 (Sincich & Horton, 2005). V1 receives strong feedback projections from V2 (Sincich & Horton, 2005), however, the feedback

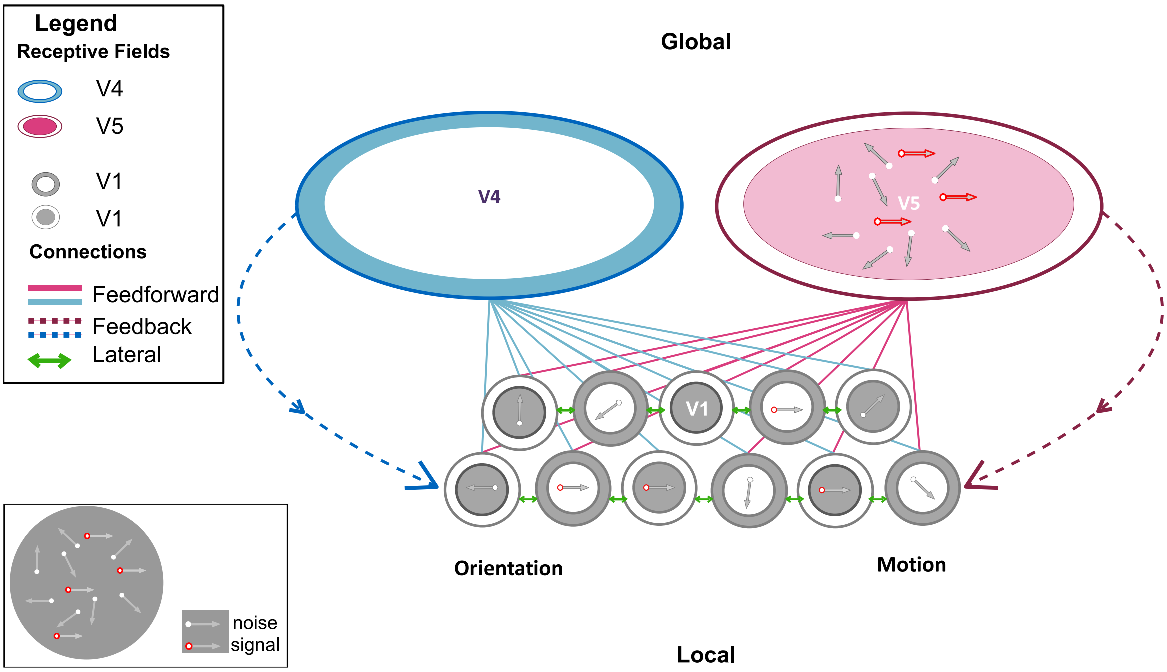


Figure. 1.5 Receptive cells in the lower levels of the visual hierarchy respond to simple features of a task within a small area of the visual field. In contrast higher levels pool sensory information across a broad number of lower level receptive cells. This information is then used to integrate the signal and segregate the noise to determine a global percept.

connections from V5 to V1 are reported to cover the most territory of all the feedback connections (Rockland & Knutson, 2000). The feedback (or re-entrant) connections terminate in different combinations within various layers throughout V1 (Rockland & Knutson, 2000; Sillito et al., 2006) and are rapidly updated (Hupé et al., 2001; Pascual-Leone & Walsh, 2001). Recordings from macaque monkeys have suggested that there is almost no delay for information processed in V5 to be fed back to lower areas, and 'feedback' from V5 is present prior to the bottom up information from the feedforward connections (Hupé et al., 2001; Pascual-Leone & Walsh, 2001). Consequentially, this may result in a predictive coding strategy when interpreting the scene, where higher level cortical areas make predictions about lower-level sensory activity (Gilbert & Li, 2013). Recent evidence supports an independent motion pathway in humans, that provides a direct link from lateral geniculate nucleus (LGN) to V5 (Bridge et al., 2015). Projections are targeted retinotopically to V1 in locations lying within the V5 receptive field (Sillito et al., 2006). Thus, as V5 responses are influenced by the bottom-up responses from V1, the top-down responses from V5 shape the responses of the V1 cells which provide much of the sensory input (Sillito et al., 2006). Transcranial Magnetic Stimulation (TMS) has been used to stimulate the re-entrant connections from V5 to V1, enhancing the perception of coherent motion (Romei et al., 2016a). Using a novel paired cortico-cortical TMS protocol (ccPAS) to induce Hebbian plasticity, observers' thresholds for motion detection were reduced when the feedback connections from V5 to V1 were stimulated. However, the improvement in perception was critically dependent on the timing and direction of stimulation. There was no change when the feedforward connections from V1 to V5 were stimulated (Romei et al., 2016b). This also suggests that these re-entrant connections are malleable (Romei et al., 2016b). Using a similar method, a direction-selective improvement was induced by pairing subthreshold stimulation with the simultaneous presentation of direction-specific moving stimuli (Chiappini et al., 2018). This provides additional support for the accumulating evidence that the backward projections from direction-tuned neurons in areas such as V5 play a role in the perception of global motion coherence tasks.

1.4 Integration and Segmentation

While the visual system is comprised of specialised functional areas and processes sensory information in a hierarchical manner, our view of the world is not segregated by colour, depth, shape and form. Rather, we see an integrated and unified world (Livingstone & Hubel, 1988). In the early visual areas, a neuron's receptive field is relatively small (see Figure 1.5), and each neuron only responds to a small region of the visual field (Gilbert, 1998; Hubel & Wiesel, 1968). At this level, the neuron will respond to certain features of the sensory input, such as the orientation or direction of motion of a feature (Hubel & Wiesel, 1959). One of the roles of V1 is to segregate visual input for local processing. However, the sensory system needs to combine the responses from several V1 receptive fields in order to integrate relevant, and segment irrelevant information, and to combine information across scale, orientation and position, and across multiple sensory streams (Braddick, 1997). Ambiguous or conflicting first-stage signals need to be integrated over a wider spatial area to provide a global representation (Adelson & Movshon, 1982; Wilson et al., 1992). Perceptually integrating local signals into a coherent organised object is a challenging task, especially in noisy environments (Liu et al., 2017).

Perception is limited by how well the visual system can segment a signal from a noisy or irrelevant background (Barlow, 1957; Levi et al., 2005). External noise is any variability in a stimulus that masks or corrupts the signal (Gold et al., 2004), making it more difficult to detect. The more external noise that is added to a stimulus, the more difficult the signal is to detect. Internal noise is reflected in the variability in detecting a signal that stems from within the visual system. Making a perceptual decision requires the brain to pool and combine multiple sources of neural activity, in order to make a correct decision (Yang & Shadlen, 2007). Internal noise is the irrelevant neural firing that is not correlated with the presence of the target signal (Webster et al., 2011). The sensory system needs to combine the segregated V1 inputs, segment the signal and the noise, and finally integrate the signal to create an unambiguous global percept.

1.4.1 Motion Perception

A well studied examplar of the integration and segmentation process is for the perception of motion. The perception of motion is hypothesised to occur as a two-stage process (Bex & Dakin, 2002; Braddick, 1993). At the first stage, spatial frequency- and orientation-tuned mechanisms in V1 encode the motion signals that occur locally within the receptive fields of individual neurons (Burr & Thompson, 2011; Movshon & Newsome, 1996). V1 neurons provide the basis for local motion, or the first stage of analysis (Braddick, 1993) Figure 1.5 (a). However, the sensory system needs to combine the responses from several V1 receptive fields in order to integrate multiple scales, directions and velocities of motion (Braddick, 1997).

The perception of local and global motion has been studied in humans behaviourally using psychophysical tasks. Local motion tasks describe motion tasks that provide only 1 or 2 signal dimensions (direction and speed respectively) (Edwards et al., 2013). However, this analysis is restricted to a small area of the visual field, less than 1 degree. (Smith et al., 1994). In local tasks the discrimination decision does not require the participant to integrate

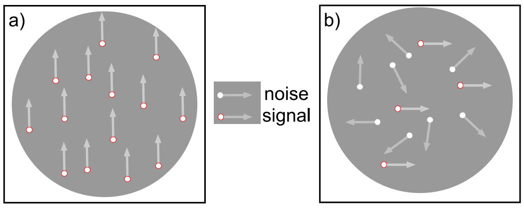


Figure. 1.6 Examples of a local and global motion coherence task. Both schematics represent white dots in motion. White dots with a red border indicate signal dots, and those without a border added noise (distractors). The arrows are representative of the direction of motion of the dots. (a) local task requires (for example) a left right decision. Dots are moving up towards the left by 1°. In a global motion task (b) a left/right response is also required, however this is obtained by integrating the signal dots and inferring an average direction.

or segment signal from noise. For example, in Figure 1.6 (a) the local task requires (for example) a left right decision. In this schematic the dots are moving up and towards the left by only 1°. This is perceptually a difficult task, but does not require global level processing, as a decision of direction can be made from tracking one dot.

Williams and Sekuler (1984) identified that when different motion vectors were combined observers perceived a coherent global motion. They proposed that the directions of the individual movements are averaged across time and space to generate the perception of a single coherent direction of motion. There are several methods of creating motion stimuli using random dot kinematograms. The most common involves a subset of dots moving coherently (signal) and others moving randomly (noise). For example in Figure 1.6 (b) the noise dots moving in random directions would make the signal dots, moving rightwards, more difficult to detect.

1.4.2 The physiology of motion perception

Since V5 is known to contain a majority of direction-sensitive neurons, Britten et al. (1993) used trained rhesus monkeys to perform a near threshold direction discrimination task using global motion stimuli while recording from this area. They found that responses varied linearly with the correlation of stimulus coherence and signal strength, and established that there was a relationship between the perception of motion, and activity in the V5.

A number of areas within the visual cortical hierarchy play a functional role in processing motion. Areas V2 and V4 have a role in processing moving orientation signals (An et al., 2012; Li et al., 2013). V3A also plays a role in several aspects of motion processing (Felleman & Van Essen, 1987), with 76% of neurons being selective for orientation and 40% showing strong direction selectivity. However, evidence from lesion studies (Newsome et al., 1989; Rudolph & Pasternak, 1999), extra-cellular recordings (Britten et al., 1992; Newsome et al., 1988) and neuroimaging in humans (Rees et al., 2000) support area V5 as a brain area that is heavily involved in processing global motion (Braddick, 1993; Britten et al., 1992; Cowey et al., 2006; Newsome et al., 1989, 1988; Rudolph & Pasternak, 1999). Most neurons in V5 are strongly direction selective (Hedges et al., 2011; Rees et al., 2000), and the evidence for the role it plays in spatially integrating motion signals is well supported by non-human primate data, and neuroimaging studies in humans (Furlan & Smith, 2016; Lui et al., 2007; Newsome et al., 1988; Pasternak & Merigan, 1994). Receptive fields in V5 can be up to tenfold larger than those in V1. V5 neurons have large, broadly tuned receptive fields that sum the responses of a set of V1 neurons across space, temporal frequency and spatial frequency (Simoncelli & Heeger, 1998). Furthermore, the feedback projections from V5 to corresponding retinotopic locations around the stimulus location of the V1 cells (Sillito et al., 2006), have been found to cover the most territory (Rockland & Knutson, 2000). This suggests that there is an important role for the feedback projections in the perception of motion.

The specific involvement of human MT for the analysis of global motion has been found using fMRI (Vaina et al., 1998). Several studies have also found perceptual learning to occur using global motion stimuli (Ball & Sekuler, 1982; Huxlin, 2008; Sundareswaran & Vaina, 1995; Vaina et al., 1995). In these cases, learning is not likely to involve those areas in V1 that have small receptive fields, and more likely to involve regions such as V5 (Vaina et al., 1995).

Whereas neurons in V1 are selective only for the component of motion that is orthogonal to their preferred spatial orientation (see Simoncelli and Heeger (1998)), the neurons in V5 combine the responses of V1 neurons. Therefore, local and global stages of motion processing differ in their tuning for orientation and scale. Psychophysical studies have shown that global motion detectors have relatively broad spatial frequency tuning (Bex & Dakin, 2002). This is consistent with single-cell recordings from area V5 in marmoset monkeys that exhibit bandpass spatial and temporal frequency tuning, with a preference for low spatial frequencies (Lui et al., 2007). Bex and Dakin (2002) undertook a comprehensive psychophysical investigation into the frequency tuning of both local and global motion detectors. Local motion detectors were found to have receptive fields that are narrowly tuned for spatial frequency. In contrast, global motion detectors were found to have much broader spatial frequency tuning, with a preference for low spatial frequencies, in line with their larger receptive fields.

Using a noise masking paradigm, Amano et al. (2009) investigated if the specificity of the local motion spatial frequency tuning was preserved once pooling for global motion has occurred. Stimuli were composed of either drifting Gabors or plaid elements that were either signal or noise. Signal elements were defined by a consistent drift speed and direction, and noise elements were defined by random speeds and directions. Frequency tuning was determined by examining how noise elements at one spatial frequency interfered with the extraction of the signal at another frequency. Amano et al. found that thresholds increased as the spatial frequency of the noise elements was reduced, and concluded that this is indicative of a broadband, low-pass tuning of motion pooling. The results obtained by Amano et al. are consistent with studies that suggest that global motion detectors pool over spatial frequencies and have a preference for low spatial frequencies (Bex & Dakin, 2002). Moreover, this is further supported by an fMRI study (Henriksson et al., 2008) which found that while areas V1,V2,V3 and V4 showed bandpass spatial frequency tuning, area V5+ exhibited an attenuated response for high spatial frequencies, but no reduction in responses for low spatial frequencies. Amano et al. (2009) proposed that, within V5, there is a “motion-pooling mechanism” that demonstrates a broadband, low-pass tuning. Pooling of visual sensory information is an important step for integration and segregation.

1.4.3 Depth Perception

Binocular depth perception depends on our ability to determine the difference in position of corresponding points between the two eyes' images. The disparity between the retinal images is used to construct a three dimensional layout of the environment (Neri, 2004; Parker, 2007). However, unlike area MT for motion processing, there is no single area of specialisation for processing depth, and binocular cells are found in almost all regions of the visual hierarchy (see Neri (2004) and Parker (2007) for discussion). Similar to the tuning properties of receptive cells for spatial frequency, V1 receptive cells are also tuned to specific retinal disparities (also referred to as *absolute disparity*) (Parker, 2007).

This first stage of processing provides a response that depends on the absolute disparity of features, the difference in location of features in each eye's image in reference to the point of fixation (see Figure 1.7). The correspondence problem is solved by using a matching measure to determine the image regions in each eye that are most similar in terms of the variation in local luminance intensity. This similarity matching can for example be based on the correlation of local intensity values (Allenmark & Read, 2011; Cormack et al., 1991; Fleet et al., 1996; Julesz, 1971). At the correct disparity offset between the left and right eyes, each point is expected to have similar luminance values in each eye. Local features at each points are matched for similarity (Fleet et al., 1996). Firing rates for neurons tuned for that disparity between the retinal images lead to high interocular correlation. Neurons not tuned to that disparity will tend not to show high correlation, and thus are expected to have reduced firing rates, since points are likely to be be unmatched in their luminance (Hibbard et al., 2016).

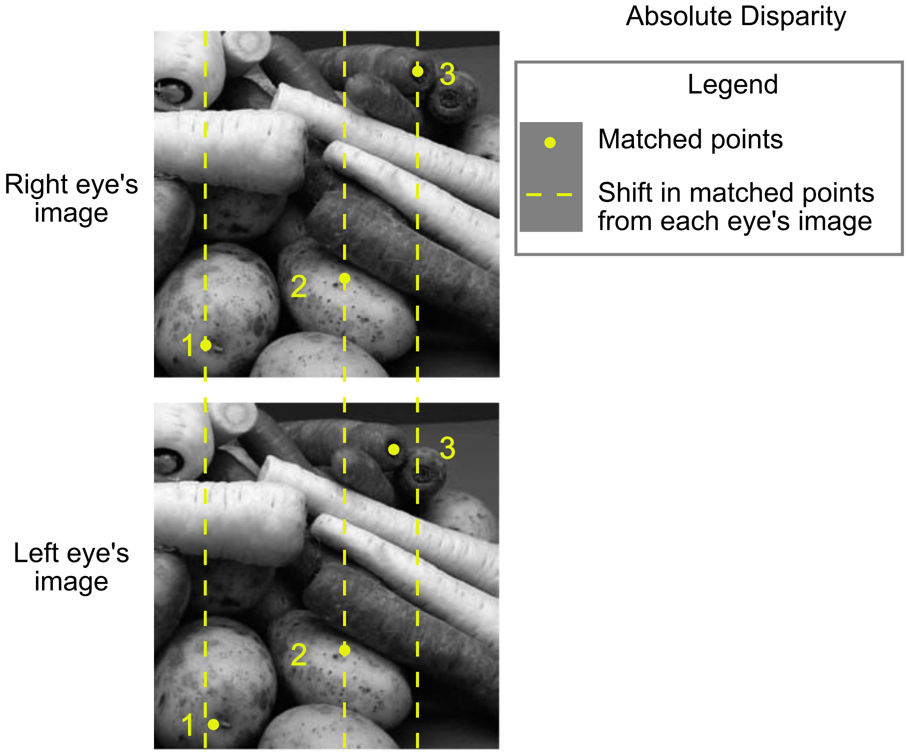


Figure. 1.7 The separation of human eyes results in a slightly different retinal image which provides an important depth cue, the disparity, or difference, obtained from each eye is matched by the visual system in order to reconcile which points in one eye's retinal image corresponds to a matching point from the other eye's image. Absolute (retinal) disparity changes with a shift in fixation, here 3 points have been marked on the right (top) and left (bottom) eye's retinal images. Each eye receives a slightly different image depending on where the eyes have fixation. For point **1**, there is a small shift rightwards in the left eye's image in comparison to point **1** in the right image. Conversely comparing the differences in the location of point **3**, there is a leftward shift of that same point in the left eye's image. Since point **2** appears to fall on the same location from each eye's image, we can assume that (i) this was the point of fixation for this image and (ii) because the locations fall on the same retinal location there is no disparity (zero disparity) at this location and therefore no depth. Whereas, point **1** will appear closer and **3** further away. Images reproduced with permission (Hibbard et al., 2014)

.

Single cell recordings in the macaque have identified that this first matching process occurs in area V1 (Cumming & Parker, 1999). However, area V1 has been found to only provide an estimate of absolute or retinal disparity, which is not sufficient to provide the perception of depth (Cumming & Parker, 1999). Human depth perception is dependent on relative disparity (see Figure 1.8 (a-b)) and the initial stage within V1 does not provide a reliable estimate of the binocular disparity required to match the individual features from the left and right retinal images (Cumming & Parker, 1997; Fleet et al., 1996). The perception of stereoscopic depth relies on a more complex process of cross-correlation, to estimate the relative disparities between individual features (Hibbard et al., 2014).

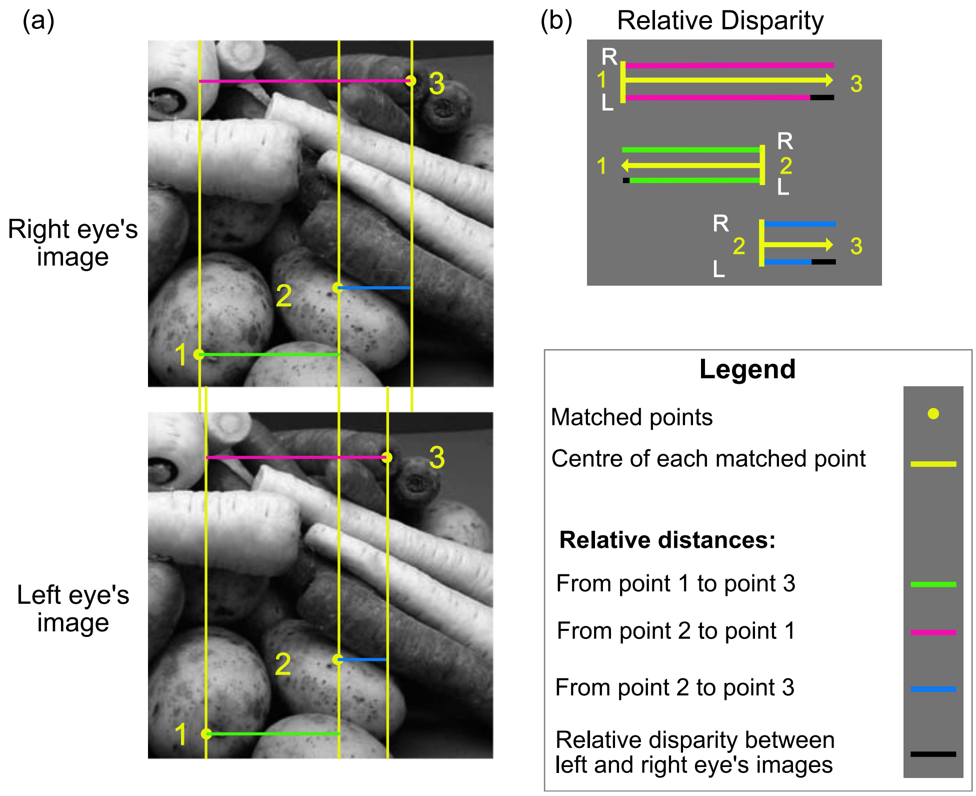


Figure. 1.8 (a) Human depth perception needs to compute relative disparity in order to provide a reliable estimate of depth. In contrast to absolute disparity (Figure 1.7) relative disparity is not contingent on the point of fixation, rather it estimates the difference in the relative distance between matched points for each eye. For example,(b) relative disparity has been estimated visually, by subtracting the measured distances between any two matched points for each eye. For example, at point **2** (also most likely the point of fixation) and there is zero disparity for the matched points in the left and right eye's retinal image. However, the difference in the distance, between the images from each eye, for matched point **2** and **3** is larger that the difference between **2** and **1**. The relative difference between each matching point of luminance contributes to the estimation of a reliable three dimensional image. Images reproduced with permission (Hibbard et al., 2014)

.

However, from the initial stage of depth processing there will be a large volume of responses in V1 that are not tuned to the correct disparity. Integration and segmentation of relevant and irrelevant responses is thus a crucial step in estimating depth (Fleet et al., 1996; Hibbard et al., 2016; Smallman & MacLeod, 1997; Tanaka & Ohzawa, 2006). Cells tuned to relative disparity are found in many visual areas, including V2 (Assee & Qian, 2007;

Bredfeldt & Cumming, 2006; Bredfeldt et al., 2009; Thomas et al., 2002) and V4 (Umeda et al., 2007) and are responsible for eliminating false matches by pooling the responses across position, orientation and scale (Fleet et al., 1996).

1.4.4 Global Form

Similar to motion and depth, the perception of local and global form has been investigated physiologically (Adab & Vogels, 2011; Gawne, 2002; Yang & Maunsell, 2004) and psychophysically (Garcia et al., 2013; Wilson & Wilkinson, 1998). The encoding of global form relies on the conversion of local information such as edges, orientation or location into a meaningful and coherent percept (Ostwald et al., 2008). Furthermore, despite the complex cross-correlation computation required to make to estimate a three dimensional world (Parker, 2007) and the monocular blind spot in each eye's retinal image (Durgin et al., 1995), the human brain is an expert shape and object recogniser (Kourtzi et al., 2005). Behavioural studies use tasks that generate a global stimulus that, like motion, needs integrating or segmenting. Area V1 extracts the local information across small areas of the visual field, such as the orientation or spatial frequency of a feature (Hubel & Wiesel, 1968). These basic components are combined to create the perception of global shape and structure with complex curves, contours and depth (Wilson & Wilkinson, 1998).

Evidence for the cortical areas involved in processing the perception of form and objects is primarily obtained using electrophysiological research in primates (Roe et al., 2012). The effects of learning on the receptive field structure of visual neurons have been measured directly using single-cell recordings in area V4 of the rhesus monkey while performing a behavioural task (Yang & Maunsell, 2004). The monkeys were trained (and rewarded with juice) to indicate whether the orientation of two interleaved counterphased Gabors were the same. The physiological responses did not show evidence for a learning effect in V1 neural locations, but did at the higher form processing area V4. Similarly, Adab and Vogels (2011) obtained single cell recordings in the Rhesus monkey during perceptual training for oriented noise masked Gabors. Findings from the study demonstrated that for the two trained orientations there was a decrease in the variability in the neural responses, and an increased difference between the neural responses for the two trained orientations. Adab and Vogels (2011) argue that despite the evidence for perceptual learning occurring as a result of a reweighting of the readout of responses at a decision stage, models of perceptual learning also need to consider modifications at the level of the sensory cortex.

Examples of stimuli used for assessing global form behaviourally are Glass patterns (Glass, 1969). Glass patterns typically consist of pairs of elements aligned to show a global form. Noise is introduced to these stimuli by varying the orientation of the elements relative to each other, or by adding noise elements at random orientations. In Glass patterns, the aim is to detect the coherent orientation. This involves potentially two processes, the integration of relevant contours (Field et al., 1993; Hess et al., 2003), and the discrimination of the elements of coherent orientation against noise elements.

1.5 Perceptual learning for visual rehabilitation

The success of perceptual learning has prompted discussion that the early vision system may be able to change. An important consequence of this plasticity, is the potential for a direct application in clinical populations, where there are deficits in the sensory cortex.

1.5.1 Visual impairments

Vision may be impaired by a range of conditions that can occur across the lifespan. Amblyopia is a common developmental impairment, affecting 2-4% of the population, that occurs if there has been an abnormal visual experience. Its severity is often dependent on the age at which the abnormality occurred (Astle et al., 2011a; Levi, 2005). Amblyopia impairs spatial vision, and is characterised by a reduction in visual acuity (how clear the visual percept is) and contrast sensitivity (the ability to detect small differences in luminance) (Levi et al., 1997). Amblyopia can be reversed if treatment is applied early in development, such as in young children, but not for older children or adults (Levi, 2005; Levi et al., 1997). Age related macular degeneration is often caused by the degeneration of the photoreceptor cells within the macula (a small area near the retina) (Plank et al., 2014) that results in an irreversible loss of central vision. Macular degeneration is the leading cause of visual impairment in developed countries for elderly populations (Liu et al., 2012b) and occurs in approximately 13% of people aged 60 or older (Liu et al., 2012b), and is known to reduce visual acuity and diminishes the perception of contrast (Liu et al., 2012b; Maniglia et al., 2016; Plank et al., 2014). A loss of visual field is one of the most common consequences of stroke, second only to weakness in the arms or legs. It is reported that 60% of stroke survivors will have a visual impairment after their stroke and, while there is a 50% chance of spontaneous recovery within the first 30 days, dropping to 20% chance between 1 – 6 months, and after 6 months, no chance of recovery is expected. The number of people who can expect a full recovery is extremely rare, less than 5% of the population (Urbanski et al., 2014).

Irrelevant of the cause, vision loss in adulthood is considered permanent. To date, there are no treatments that can fully restore lost vision. However, perceptual learning rehabilitation programmes have also been used, to treat or overcome some of the deficits for adults with amblyopia (Astle et al., 2011a,b; Hess et al., 2003; Levi & Li, 2009; Levi & Polat, 1996; Levi et al., 1997; Polat et al., 2004), myopia and presbyopia (Camilleri et al., 2014b; Polat, 2009), nystagmus (Campana & Maniglia, 2015),deficits in stereopsis (Ding & Levi, 2011), individuals with cortical blindness (Das et al., 2014; Huxlin et al., 2009; Kasten & Sabel, 1995; Nelles et al., 2001; Sahraie et al., 2006; Trevethan et al., 2012) and macular degeneration (Maniglia et al., 2016; Plank et al., 2014).

While perceptual learning offers a potentially promising route for rehabilitation, improvements are often highly specific to the features of the trained stimuli, limiting its effectiveness as a general tool (see Dosher and Lu (2017) for a comprehensive review). A deeper understanding of how behavioural training promotes neural plasticity is still needed (Maniglia & Seitz, 2018).

1.5.2 Neuro-modulation and rehabilitation

Recently, there has been interest in evaluating how the perceptual learning process can be manipulated using neuromodulation techniques such as transcranial magnetic stimulation (TMS) or transcranial electrical stimulation (tES). These techniques are proposed to accelerate neural plasticity by either increasing the activity of the sodium ion channels, or by modulating the firing rate of neurons being stimulated respectively (Camilleri et al., 2014a). Transcranial random noise stimulation (tRNS), a type of electrical stimulation that alternates current at random frequencies, has been shown to improve contrast sensitivity in individuals with amblyopia, myopia and presbyopia (Camilleri et al., 2014a). Improvements were only found in those individuals who received both tRNS and behavioural training (perceptual learning). This study suggests that it is possible to reduce the long training times usually involved with perceptual learning regimes, and obtain similar results within a much shorter period.

1.5.3 The debate against cortical reorganisation

The proposal that early visual areas are malleable is still controversial, and the nature of plasticity within the visual system is argued to be uncertain, inconsistent and plagued with unanswered questions (Wandell & Smirnakis, 2009). While there is no doubt that neural plasticity occurs, some cortical networks may require more stability, and re-routing paths in the visual system may have a negative impact, disrupting computational circuits in correctly functioning perceptual processes (Wandell & Smirnakis, 2009). Some systems remain plastic during the course of the lifetime such as those used in the processing of language (Neville

& Bavelier, 1998), memory (Neves et al., 2008) and motor systems (Doyon & Benali, 2005). It is argued that sensory systems are more susceptible to adaptation than to cortical reorganisation (Wandell & Smirnakis, 2009). Adaptation is a short-term change that is made in response to a change in the sensory environment. Adaptation occurs very quickly and is known to produce several established perceptual effects, including the motion after-effect (Barlow & Foldiak, 1989), however unlike perceptual learning these are not persistent, and do not appear to rely on structural changes within the visual cortex (Wandell & Smirnakis, 2009). Adaptation is thought to occur as a result of interactions and mutual inhibition (Barlow & Foldiak, 1989; Movshon & Lennie, 1979).

Wandell and Smirnakis (2009) argue that certain areas within the brain require stability. For example, disrupting stereo perception during the critical period has been shown to result in irreparable damage (Hubel & Wiesel, 1970). These neural pathways are established through early visual experience by coordinating and combining the separate inputs from each eye (Parker, 2007; Seitz, 2017). To extract these differences, binocular neurons match the individual features from each eye's image to create stereo perception (Parker, 2007). In order to compute this, the perceptual and sensory systems need to be flexible enough to allow for growth during adolescence but robust enough to efficiently provide an accurate sensory experience (Levi, 2005). The binocular connections are known to be highly vulnerable during the first 18 months for children (Banks et al., 1975; Levi, 2005), and the most dramatic increase in interpupillary distance occurs within the first 12 months of life, although this continues to change into early adulthood (Fesharaki et al., 2012).

Smirnakis et al. (2005) used fMRI (functional magnetic resonance imaging) and microelectrode recordings to detect changes in the macaque early visual cortex after binocular retinal lesions. Their results found no evidence of cortical re-organisation near the border of the lesion zone in V1 of adult macaque, in direct contrast to the findings of previous microelectrode studies that found evidence of plasticity (Gilbert & Wiesel, 1992; Heinen & Skavenski, 1991; Smirnakis et al., 2005). Using a combined methodology of microelectrode and fMRI provides the advantage of sampling the whole brain, and unlike single cell recording avoids the bias of overestimation from a single neuron (Sereno, 2005). Similarly, Baseler et al. (2011) evaluated cortical maps from adult humans with macular degeneration and juvenile macular degeneration (acquired as juvenile), finding no evidence of remapping of the early visual areas in adults, and no dependency was found for the age at which retinal damage occurred.

Haak et al. (2015) proposed that while retinal lesions may cause rapid adaptive changes in the retinotopic receptive fields within the zone of a lesion, the mismatch between the difference in stimulus locations within the visual field may result in confused signals within the subsequent layers of processing.

Despite the cautions about the degree of plasticity in the adult visual cortex, continued interest in perceptual learning and its benefits to individuals with cortical visual deficits continues to be of interest to researchers (Das & Huxlin, 2010; Hadid & Lepore, 2017; Huxlin et al., 2009; Urbanski et al., 2014). While there is conflicting evidence for the susceptibility of the adult primary visual cortex for cortical remapping, it is possible that any changes, such as those from perceptual learning, may occur later in the visual hierarchy (Ahissar &

Hochstein, 2004; Haak et al., 2015; Hochstein & Ahissar, 2002). Therefore an ongoing theoretical question in perceptual learning concerns the location of the learning mechanism within the processing stream.

1.6 Mechanisms and Models of Perceptual Learning

An important theoretical question in perceptual learning concerns the location in the processing stream of the learning mechanism(s). To dissociate between models and mechanisms of perceptual learning, we use three general categories i) the How ii) the What and iii) Where. Models of perceptual learning aim to resolve *how* learning occurs. Understanding which mechanisms underlie this learning address the *what* and *where* (and possibly *why*) questions. In this way we can consider a mechanism (the *what*) and an attribute or characteristic of the learning model (Zednik, 2018).

1.6.1 What, Where Why? Mechanisms of Perceptual Learning

In order to begin to unravel the *what*, *where* and *why* questions, the behaviours of the system in question needs to be understood, and the sensory output mapped to the behavioural output (Zednik, 2018). In this section we will consider some of the characteristics and behaviours associated with the learning process.

Specificity

A hallmark of perceptual learning is its specificity to the features of a stimulus. The specificity of learning has been a philosophical discussion point since the early 19th century. Classical thinkers believed that studying any field would act as “mental exercise" to increase the brain’s ability and capacity for learning in general (Gilbert et al., 2001). However, this was challenged by Edward Thorndike who proposed that learning was specific to the discipline (Gilbert et al., 2001). As such, specificity describes the phenomenon where improvement after training is highly restricted to the stimulus features of the trained task (Dosher & Lu, 2017). Studies have identified that learning is specific for orientation (Fahle, 1997; Fiorentini & Berardi, 1980; Schoups et al., 1995), spatial frequency (Fahle, 2005; Fiorentini & Berardi, 1980; Sowden et al., 2002), direction of motion (Ball & Sekuler, 1982), hyperacuity (Fahle & Edelman, 1993; Levi et al., 1997; Saarinen & Levi, 1995; Weiss et al., 1993), the retinal location of stimulus features (Fahle & Edelman, 1993; Fahle et al., 1995; Karni & Sagi, 1991; Poggio et al., 1991; Schoups et al., 1995; Shiu & Pashler, 1992) and the eye to which they are presented (Fahle et al., 1995; Karni & Sagi, 1991; Poggio et al., 1991). For example, in one study learning did not transfer after changing the orientation of a trained stimulus (Fiorentini & Berardi, 1980), nor when moving a texture discrimination stimulus as little as 3° from the trained position (Karni & Sagi, 1991). Similarly, observers were unable to transfer an increased performance in a trained Vernier acuity task to another target that had been rotated by 90° (Fahle & Edelman, 1993) or when using an untrained eye (Herzog & Fahle, 1997). Training to discriminate the direction of motion also does not transfer to untrained locations and is specific to the direction of motion of the trained stimulus (Ball & Sekuler, 1982, 1987). As a result of the specificity to low-level visual features of the stimuli, learning was often explained in terms of a change at the level of the neurons that code for the specific stimulus (Gilbert et al., 2001). Early psychophysical research investigated learning and transfer using lower-level visual tasks that match the receptive fields of neurons in the lower visual processing areas (Karni & Sagi, 1991). At the lowest level, V1 and V2 represent simple visual dimensions, such as the position, orientation and scale of local image features (Ahissar & Hochstein, 1997). The receptive fields of cells in V1 are small and only respond to a very restricted area of the visual field (Furlan & Smith, 2016; Lamme, 2003; Simoncelli & Heeger, 1998) and are tuned to orientation (DeValois et al., 1982; Hubel & Wiesel, 1962), spatial frequency (DeValois et al., 1982; Movshon et al., 1978) and those that are motion-specific are tuned for direction (Livingstone & Hubel, 1988; Movshon & Newsome, 1996). Thus, learning that is restricted to the type of task or features of training is said to be specific.

Transfer

Dosher and Lu (2017) proposed considering specificity as a 'graded phenomenon', whereby there is some specificity and some transfer that can depend on several factors, including the level of processing in the visual hierarchy (Fine & Jacobs, 2002) the training procedure used (Hung & Seitz, 2014; Xiao et al., 2008) and the extent of training (Sowden et al., 2002). There have, however, been studies where some transfer of learning has been found, in tasks such as orientation (Ahissar & Hochstein, 1997; Jeter et al., 2009; McGovern et al., 2012) and motion (Huang et al., 2007; Levi et al., 2015; Liu, 1999) and across retinal location (Xiao et al., 2008). For example, while motion discrimination was found to be specific to the trained location and direction, there was limited inter-ocular transfer (Ball & Sekuler, 1982, 1987). This hints to the level of processing (for that task) occurring earlier in the visual hierarchy, prior to binocular integration (Gold et al., 2004). Early psychophysical studies found little perceptual learning for luminance contrast detection (Adini et al., 2002). More recently, learning has been found to occur as long as extensive training is provided (Sowden et al., 2002; Yu et al., 2004). The improvement in contrast sensitivity has been found to be highly tuned to spatial frequency, specific for the trained eye and retinal location of training, but not selective for orientation (Adini et al., 2004; Kuai et al., 2005; Sowden et al., 2002; Xiao et al., 2008; Yu et al., 2004; Zhang et al., 2010, 2008).

Several explanations have been put forward for when transfer may occur. In support of Wandell and Smirnakis (2009), dummy trials were included within an oriented texture discrimination task, in order to disrupt sensory adaptation (Harris et al., 2012). In this study, it was found that learning generalised to new locations. Since perceptual learning occurs as a result of repeated and extensive exposure, specificity may be the results of adaptation to tasks require pattern discrimination, which may induce changes in the local sensory processing areas (Harris et al., 2012). A theoretical model of transfer is the Reverse Hierarchy Theory (Hochstein & Ahissar, 2002). This model proposes that transfer is a top-down process, and the degree of transfer is dependent on the characteristics of the receptive fields involved in performing the training task. Transfer occurs as a result of modification of neurons found in higher cortical levels, where receptive fields generalise over image location, spatial scale and other image properties. In contrast, specificity occurs at the lower cortical levels where the receptive fields are localised (Hochstein & Ahissar, 2002).

An early proposal was that transfer may depend on the difficulty of the trained task (Ahissar & Hochstein, 1997). This was based on the results of a feature detection task that varied the degree of gradient of a target in comparison to distractors. Either easy (30°) or difficult (16°) tasks were presented at varying stimulus onset asynchronies (SOAs). Performance was higher for targets with longer SOAs and a greater tilt, and lower for shorter SOAs and smaller tilts. The results indicated that the group who were trained on the easier 30° task transferred improvement to untrained orientations and locations. Conversely, those trained on the difficult task, with a 16° tilt, showed specificity for the trained location and orientation (Ahissar & Hochstein, 1997). Similarly, transfer of learning across tasks was tested using tasks in which stimuli differed in terms of orientation (clockwise or anti-clockwise direction), curvature (which of two was more curved), or global form (which of two contained a global structure) (McGovern et al., 2012). As predicted, all groups improved substantially on their trained task. However, transfer occurred in line with the complexity of the trained and transfer tasks. When observers trained on orientation (a simple task), they improved on curvature (an intermediate task), but not global form (a complex task). Those who trained on global form showed an improvement on the curvature task, but less of an improvement on the orientation tasks. Those who trained on curvature did not improve on the global form task or the orientation task. McGovern et al. (2012) proposed that their findings link to models where plasticity occurs over different visual areas across the visual hierarchy. This is consistent with the Reverse Hierarchy Theory that accounts for transfer to less complicated features, but not to those more complicated.

For motion, transfer occurred asymmetrically from a discrimination task to a detection task, but not vice versa (Huang et al., 2007). Using the same stimuli, half the observers were trained to make a direction discrimination judgement (either more clockwise or more anticlockwise than a point of reference of the extended arm of an oriented fixation cross). The other half were trained to detect which one of two presented targets contained a coherent motion signal. Participants who were trained to detect if there was coherent motion improved on their own task, but did not improve on the discrimination task. However, those who trained on the discrimination task improved on both tasks. Huang et al. (2007) argue that the exact discrimination of motion direction was irrelevant for the detection task, but both detection and discrimination of the motion signal were necessary for the discrimination task. They proposed that transfer occurs from complex to simple tasks, but not from simple to complex tasks.

It has recently been argued that transfer may be dependent on task precision rather than difficulty (Jeter et al., 2009). Specifically, precision was defined as the extent of the difference in the task. Jeter et al. (2009) manipulated the precision of the trained and the untrained tasks but kept task difficulty constant by monitoring performance accuracy. In an orientation discrimination task (using a staircase procedure), it was demonstrated that improvements were specific to the trained elements for conditions with high precision, but not low precision. Low precision conditions had an orientation difference of +/- 12° and high precision conditions had an orientation difference of +/- 5°. Their results suggested that transfer depended on the precision of the transfer task, and not the training task. Transfer was observed for low-precision transfer tasks only (Jeter et al., 2009).

A number of studies have demonstrated transfer following a double-training protocol (Hung & Seitz, 2014; Xiao et al., 2008; Zhang et al., 2010). Double-training involves overt training on a task-relevant feature in addition to exposure to a task-irrelevant feature

(Watanabe & Sasaki, 2015). Task-irrelevant features do not contain information that is required to complete the task (Seitz & Watanabe, 2009). While learning from exposure alone is not sufficient to give rise to improved performance (Garcia et al., 2013; Herzog & Fahle, 1997; Seitz et al., 2006; Shiu & Pashler, 1992), using training tasks that have an embedded irrelevant feature is argued to trigger improvements in the readouts for both features (Solgi et al., 2013). Learning can be considered as a bi-directional process, whereby there is first a strengthening of the neural circuitry between the sensory areas and higher cortical areas and an increased representation of neuronal recruitment in the sensory cortex. Secondly, there is a top down transmission from higher areas (such as motor or decision areas) that descend to lower sensory levels (Solgi et al., 2013). The mechanism, allowing transfer, is argued to occur in double training, defined by Solgi et al. (2013) as ‘gated self-organisation’, referred to more colloquially as “piggybacking" (Wang et al., 2014). For example, contrast sensitivity, historically found to be specific to the retinal location of training, was found to transfer between retinal locations when observers were trained on two simultaneous tasks (Xiao et al., 2008). Observers were required to detect a relevant feature (contrast discrimination) at one retinal location, followed by detection of an irrelevant feature (orientation) at another location, and a full transfer of contrast increment sensitivity to the untrained location was found (Xiao et al., 2008). Similarly, observers who were trained to discriminate orientation, and simultaneously exposed to a second orientation, experienced a full transfer of improvement in orientation sensitivity as long as training preceded exposure, rather than being simultaneously presented (Zhang et al., 2010). The authors proposed a rule-based learning model to account for the transfer, where higher level decision units learn and re-weight the V1 inputs. They proposed that the absence of functional connections to the untrained orientation or location prevents any potential reweighting. Double training, through exposure, activates the functional connections at the new locations or orientations, enabling transfer (Zhang et al., 2010). This suggests that there may not be a straightforward correspondence between stimulus features and the neural loci involved in performing the task. Double-training has been a successful paradigm in obtaining evidence of transfer (Cong et al., 2016; Hung & Seitz, 2014; Wang et al., 2014; Xiao et al., 2008), however task difficulty during training, and observer confidence still play a vital role in the learning and transfer process (Asher & Hibbard, 2018; Asher et al., 2018b; Talluri et al., 2015).

Learning as a change to the receptive fields

One way perceptual learning could occur is exploiting the correlations between neural responses, via anti-Hebbian or Hebbian learning (Gilbert et al., 2001). Anti-Hebbian learning has the aim of reducing the correlations between the responses of neurons, so as to produce an efficient encoding of sensory information (Barlow & Foldiak, 1989; Simoncelli, 2003; Wei & Stocker, 2012). Conversely, Hebbian learning, by accentuating the correlations between responses, might be beneficial in tasks such as global motion and form, in which it is important to segregate signal from background noise (Gilbert et al., 2001).

Internal and External noise

Learning may be explained as an improved ability to filter out external noise, or a reduction in internal noise (Dosher & Lu, 1998). If the pooling of sensory information across cells reduces internal noise this would improve efficiency and learning would occur (Gu et al., 2011). Noise levels were examined in macaques, by recording from individual pairs of neurons. However, while noise correlations were weaker for those trained on a perceptual discrimination task than those who were untrained, their results also showed that optimally decoding all neurons did not predict a meaningful change in learning performance (Gu et al., 2011). The reduction in noise was considered as an an unlikely explanation for the effects of training (Gu et al., 2011). Conversely, a psychophysical study of perceptual learning for contrast sensitivity has suggested that learning occurs through a combination of filtering out external noise, and reducing additive internal noise (Solomon & Tyl, 2017).

1.6.2 How? Models of Perceptual Learning

The *how* is often the primary focus for research in perceptual learning. Models of perceptual learning play an important role in testing the complex and often conflicting results from empirical and physiological research (Dosher & Lu, 2017). There are, broadly speaking, two predominant models of perceptual learning. The first explains perceptual learning in terms of a change in the neurons that code for that feature (Adini et al., 2002; Gilbert et al., 2001). The second position argues that learning is a result of the change in weights of readout between the sensory representation and the decision units (Dosher et al., 2013; Dosher & Lu, 1998, 2009; Herzog & Fahle, 1999; Poggio et al., 1991).

A physical change at the neural level

The first position developed as a result of the evidence that improvements were highly specific for the features of the stimuli used for training for orientation (Fahle, 1997; Fiorentini & Berardi, 1980; Schoups et al., 1995), spatial frequency (Fahle, 2005; Fiorentini & Berardi, 1980; Sowden et al., 2002), direction of motion (Ball & Sekuler, 1982), retinal location (Fahle & Edelman, 1993; Fahle et al., 1995; Karni & Sagi, 1991; Poggio et al., 1991; Schoups et al., 1995; Shiu & Pashler, 1992) of stimuli and the eye to which they are presented (Fahle et al., 1995; Karni & Sagi, 1991; Poggio et al., 1991). On the basis of this feature-specificity it has been argued that the underlying brain area responsible for the learning process must be within the primary visual cortex (Karni & Sagi, 1991) where the receptive fields of cells display a high degree of specificity. The implication of this position is a potential degree of plasticity in individual neurons within an area previously thought to be incapable of structural change (Gilbert & Li, 2013; Gilbert et al., 2001; Karni & Sagi, 1991). The effects of learning on the receptive field structure of visual neurons have been measured directly using single-cell recordings. Changes in orientation and spatial frequency tuning in V1 (Ren et al., 2016; Schoups et al., 2001) and V4 (Yang & Maunsell, 2004), and increased sensitivity to motion coherence in V5 (Zohary et al., 1994), have all been found as a result of perceptual learning. Neuroimaging in humans observers has indicated a reorganisation of the visual cortex after training (Furmanski et al., 2014; Jehee et al., 2012; Shibata et al., 2011).

While the failure of perceptual learning to transfer across retinal location, orientation and other stimulus dimensions may implicate V1, learning is far more complex than can be explained by the stimulus driven, bottom-up, visual representation model (Mollon & Danilova, 1996). The changes to the tuning of individual sensory neurons are therefore not sufficient to account for the behavioural improvements achieved through training. The effects of learning on populations of neurons need to be taken into account including the correlation and decorrelation in noise between neural responses (Gu et al., 2011; Mollon & Danilova, 1996). Furthermore, a model of learning also needs to include top-down factors such as attention (see (Gilbert & Li, 2013) for a review), confidence in decision making (Ahissar & Hochstein, 2004; Talluri et al., 2015) and knowledge and understanding of the task (Herzog & Fahle, 1998).

A change in reweighting of decision units

The second position poses that the sensory representations remain fixed, and that learning occurs as a result of optimising the read-out from sensory areas (Dosher & Lu, 1998; Mollon & Danilova, 1996). Repeated experience would strengthen the connections between lower visual areas and the high level areas where the perceptual decision is then made (Petrov et al., 2006a). Rather than a physical change in the sensory neuron, learning represents a change in the weights of connections between the input, sensory encoding layer and the classification layer at which the decision is made (Dosher et al., 2013; Dosher & Lu, 1998, 2009; Herzog & Fahle, 1999; Poggio et al., 1991). Perceptual learning models that explain learning as a change in decision weights are computational models (Dosher & Lu, 2017). Like artificial neural networks, these use learning rules to deterimine the change in weights between units across a predefined network.

Summarising the *how* of perceptual learning

Behavioural studies of perceptual learning are unable to distinguish between learning as a result of changes in sensory encoding, or as a result of a reweighting of the connections between sensory and decision-making stages of processing. However, learning effects on the structure of receptive fields of visual neurons have been measured using single cell recordings. Changes in orientation and spatial frequency tuning in V1 (Ren et al., 2016; Schoups et al., 2001) and V4 (Yang & Maunsell, 2004), and increased sensitivity to motion coherence in V5 (Zohary et al., 1994), have all been found as a result of perceptual learning. In addition, neuroimaging in human observers has indicated a reorganisation of the visual cortex after training (Furmanski et al., 2014; Jehee et al., 2012; Shibata et al., 2011). However, this in itself cannot dissociate between learning models (Herzog & Fahle, 1998; Mollon & Danilova, 1996; Petrov et al., 2005). It is generally judged that physiological models are not sufficient to explain the improvements in task performance exclusively (Dosher & Lu, 2009; Gu et al., 2011).

In summary, psychophysical evidence points to perceptual learning taking place at an early stage of stimulus processing (Fahle, 2002; Gilbert et al., 2001). Neural network models point to perceptual learning occurring as a result of computational change in weightings (Dosher et al., 2013; Dosher & Lu, 1998, 2009; Herzog & Fahle, 1998, 1999; Petrov et al., 2006a; Poggio et al., 1991). Finally physiological evidence suggests that there are observable changes in neural representation in the areas responsible for learning perceptual tasks (Adab & Vogels, 2011; Britten et al., 1992; Newsome et al., 1989, 1988; Rudolph & Pasternak, 1999; Yang & Maunsell, 2004). There are a broad network of neurons involved in perceptual learning. For example a simple low-level feature task will be be influenced in the first instance by a bottom-up stimulus driven processes that may involve change(s) to the sensory representation of the stimulus (Gilbert et al., 2001). However, in addition there are also complex decision rules that change neural weightings throughout the neural hierarchy (Petrov et al., 2006b). These weightings are also influenced by top-down processes such as the level of confidence, internal or external reinforcement, attention to the task (Maniglia et al., 2018)

1.7 Feedback in perceptual learning

Despite the complicated assessment of an ever changing environment, observers can learn to interpret incoming sensory data in novel environments (Scarfe & Glennerster, 2014). Performance feedback is information that notifies a learner about their performance, and can be generated internally or be provided by an external source (Fahle, 2002). Using performance feedback allows an organism to quickly establish the appropriateness of their learning strategy, and thus, should be an advantage for those organisms who were able to do so efficiently (Shibata et al., 2009). In psychophysical experiments this often takes the form of an auditory sound that is presented after a trial to indicate a correct or incorrect response. Behavioural perceptual learning studies have shown that using external feedback can improve learning and increase efficiency (Herzog & Fahle, 1999).

The role of feedback in perceptual learning is still unclear. Several studies have found that feedback on performance has increased learning (Herzog & Fahle, 1997), or is a necessary factor (Seitz et al., 2006), but there are also those that find learning occurs without external feedback (Liu et al., 2012a; Petrov et al., 2005; Vaina et al., 1998). This results in a complicated pattern of empirical findings. However, feedback is not usually the primary area of interest in perceptual learning studies, and studies that fail to find significant results are less likely to have been published (Seitz et al., 2006).

In an early attempt to clarify the role of feedback in learning, Herzog and Fahle (1997) tested several models using a Vernier acuity task. The first group received trial-by-trial feedback, and all observers bar one displayed significant improvement. A second group trained without feedback, and provided a mixture of results, including a non-significant improvement, no improvement and some worsening. A third block of observers were provided with block feedback, in terms of a percentage correct after each block. Improvements were similar to those found for individuals receiving trial-by-trial feedback. Since feedback was not attached to an individual stimulus, they argue this type of feedback could not act as a teaching signal. A further group were provided with random trial-by-trial feedback. There was no learning and the authors concluded that uncorrelated feedback prevents learning. A group of observers receiving partial feedback, on 50% of their incorrect responses, did improve, but less than those with full feedback. Additionally, in a group receiving reverse feedback, all but one observer adapted. Herzog and Fahle (1997) concluded that correct feedback improves both the speed of learning and overall improvement in performance. In the no feedback condition, results were highly variable among observers, and, on average, no learning was found without feedback. Feedback played a role in reducing the variation in responses over observers. They suggest that external feedback has an important role to play in learning that cannot be explained as a teaching signal, since block feedback had no signal to individual stimuli, yet learning occurred at the same rate as trial by trial feedback. Furthermore, they note that learning with feedback was significantly easier than without it, and there is a positive effect of feedback, especially since manipulated feedback prevented learning. Finally, they note that learning, exclusively as a result of exposure to a stimulus, is implausible.

However, Shibata et al. (2009) reported results that contradict Herzog & Fahle's conclusions. In a total of nine groups they manipulated the accuracy of feedback for six of the nine groups. Manipulated groups received predefined feedback irrelevant of their actual accuracy. For the three remaining groups; one group had unmanipulated, accurate feedback, one group no feedback, and one group had no feedback, but were asked to judge their own level of performance. The task was a same or different discrimination task, of complex gratings. Shibata et al. (2009) defined a learning regression line for the average accuracy across subjects in the unmanipulated accurate feedback, condition, manipulated feedback was constructed by manipulating the gradient of the regression line, where participants' predefined feedback would either be the same, larger or smaller than the original. Their results revealed that learning occurred in the group that had no feedback, which was no different to performance in the unmanipulated accurate feedback condition. Observers who had manipulated feedback with a larger gradient, were facilitated in comparison to the unmanipulated accurate feedback group. Observers who received smaller gradient feedback improved at the same level as the genuine group.

Vaina et al. (1995) differentiate between perceptual learning over several days, and 'rapid' learning, which stabilises within the first 200 repetitions. Observers tested daily on a global motion task, with a signal of 25% for three days without feedback, and repeated the experiment again 10 days later. They found learning occurred extremely quickly within session for a global motion task, where observers went from chance performance to 100% correct across the 10 days. However, this improvement was lost when direction was changed from left/right to up/down, suggesting that improvements were specific to trained direction.

Moving away from rapid learning, Petrov et al. (2006a) suggested that feedback may be useful when the stimulus is difficult to detect or discriminate, where it may increase confidence and make learning more efficient. Liu et al. (2012a) predicted that there was an interaction between accuracy and feedback. Where accuracy is high for a sufficient number of trials, Hebbian learning predicts a high chance of learning, however, when accuracy is low Hebbian learning alone is erratic. Alternatively, when feedback (trial-by-trial) is provided, there should be less reliance on performance accuracy. They tested this prediction in a 6 day contrast sensitivity (with noise) paradigm using a staircase method. Observers were divided into high and low accuracy training groups, half of whom received trial-by-trial feedback, and half of whom did not. They found an interaction between feedback and accuracy; when accuracy was high, external feedback was not critical, but it was crucial when accuracy was low. Furthermore, Liu et al. (2012a) replicated the study, finding that by mixing high and low accuracy trials learning also occurred without the need for external feedback.

Seitz et al. (2006) also investigated if including easy exemplars could foster perceptual learning in the absence of feedback. Two studies investigated firstly a motion direction discrimination task with low contrast dots, and secondly an orientation discrimination (masked with noise) task using off cardinal (obliquely) oriented bars. While both groups receiving external reinforcement displayed perceptual learning effects, those that experienced no feedback failed to show learning. They concluded that internal reinforcment was not enough to generate reinforcement signals.

Since the role of feedback is not a routinely tested paradigm, with only a handful of studies that explicitly test for differences in learning as a result of feedback (Seitz et al., 2006), the role of external feedback in perceptual learning remains unclear (Seitz et al., 2006, p.972).

1.8 Chapter Outline

This thesis considers individual differences in perception, and the mechanisms of perceptual learning, in order to address a number of the topics addressed in the discussion so far. The focus of each chapter is summarised below.

- **Chapter 2** Behavioural measures to assess perceptual learning are primarily obtained through psychophysical experiments. Standard methods are often not well-suited to capturing individual differences. This chapter explores statistical methods which are better suited to studying the differences in performance, between individual and across time, which characterise performance in the psychophysical studies in this thesis.

- **Chapter 3** Visual processing is known to differ in people with migraine (O'Hare & Hibbard, 2016). This chapter measures differences in migraine contrast sensitivity and the lateral interactions of the type that underlie global form perception.

- **Chapter 4** Global tasks involve local and global levels of processing, and both feedforward and reentrant connections between the two levels. This chapter assesses the spatial frequency tuning of the transfer of perceptual learning of global motion, to both global motion and contrast sensitivity tasks, in order to understand the nature of perceptual learning for global motion.

- **Chapter 5** This chapter assesses perceptual learning in local and global form tasks, using adaptive staircase and method of constant stimuli psychophysical methods with and without feedback. This provides important insights into the role of external feedback and internal confidence in learning at local and global levels.

- **Chapter 6** Binocular depth perception is based on a process of binocular cross-correlation. This chapter uses anti-correlated random dot stereograms to understand the pooling of information across scale and between first- and second-order mechanisms in this process. Individual differences in binocular depth perception, and the relationship between performance and confidence, were also assessed.

- **Chapter 7** The final chapter summarises the findings and their contributions to understanding individual differences and learning in perceptual tasks.

Chapter 2

Inclusive psychophysics: embracing variability in psychophysical performance

2.1 Introduction

Psychophysics is both a methodology and a research area (Kingdom & Prins, 2010), and has many applications. A major focus of psychophysics is to measure and understand the sensitivity of the perceptual systems of human and non-human animals. Sensitivity describes how intense a sensory event needs to be in order for it to be detected or discriminated and can be used to identify the limits of and differences between perceptual systems (Lu et al., 2014). There are many important practical applications of psychophysics in understanding the human sensory experience. For example, studies from the United States have shown that accidents from medical emergencies occur most often at an intersection between an emergency vehicle and a motorist (see Howard et al. (2011)). As early as the 1980's, researchers have questioned the motorist's ability to detect and localise ambulance sirens under normal driving conditions (Caelli & Porter, 1980). Catchpole and Mckeown (2007) used acoustic psychophysics to evaluate different ambulance sirens and found that appropriate evasive action was dependent on the motorist's ability to localise the source of the siren. Via psychophysical experimentation, the amplitude and frequency range of the siren that were best able to support accurate localisation of the emergency vehicle could be established (Catchpole & Mckeown, 2007; Howard et al., 2011). Based on these results, it was recommended that a broad frequency range, including both low (Howard et al., 2011) and high frequency (Catchpole & Mckeown, 2007) components, should be included in warning signals. More recently, psychophysics has explored sensory experiences in Virtual Reality (VR), allowing researchers to create immersive virtual experiences by optimising the multisensory sensory information that is presented, and minimising artefacts such as lags in motion tracking and display update to acceptable levels (Scarfe & Glennerster, 2015). Optimising VR in this way allows for its successful deployment in surgery, dentistry, exposure therapy to treat phobias (García-Palacios et al., 2002), and in education and training, as well as applications in gaming and other creative industries.

Another crucial focus of psychophysics is analytical in nature, and allows researchers to make inferences about the underlying neural mechanisms of perception (Gescheider, 1997). An important extension of this body of knowledge is the use of visual psychophysics in clinical assessments of visual function including acuity, perimetery, depth and colour

(Lu et al., 2014). Knowledge about the variation in human sensory sensitivity can be used to identify differences in neural mechanisms for clinical populations such as those with migraine (Asher et al., 2018a; O'Hare & Hibbard, 2016), Myalgic Encephalomyelitis (M.E.) (Ahmed et al., 2018), autism spectrum disorders (Manning et al., 2015; Simmons et al., 2009), schizophrenia (Shergill et al., 2005) and bipolar disorder (Krug et al., 2008). Psychophysical tasks have also been used to help those with reduced or damaged vision, such as amblyopia (Levi & Li, 2009), macular degeneration (Plank et al., 2014) and stroke (Campana & Maniglia, 2015; Huxlin, 2008). Since psychophysical methods do not rely on introspective, verbal report methods, they can be used with young infants (Braddick et al., 2003), children (Manning et al., 2018) and non-human participants such as monkeys (Law & Gold, 2008), cats (Movshon et al., 1978; Muckli et al., 2005) and even praying mantises (Nityananda et al., 2018).

2.2 Psychophysical methods

Psychophysical experiments behaviourally investigate the relationship between the physical dimensions of stimuli and subjective experience. This allows experimenters to measure aspects of perception such as the sensitivity to the presence of a stimulus, or to differences between stimuli, and biases when making judgements of the physical properties of objects. Experiments typically involve the methodical manipulation of one or more dimensions of a stimulus, such as its intensity or presentation time, and requiring participants to make a prescribed response. For example, they may be asked to indicate whether a stimulus was present or absent, or which of two was brighter. Participants will differ in the signal strength they will require in order to reliably detect or discriminate the stimulus. In order to identify how sensitive each participant is to the stimulus, the researcher would typically present numerous targets at varying levels of intensity. A stimulus with a higher intensity (or intensity difference) has a better chance of detection (or correct identification). Figure 4.1 shows a commonly used stimulus in vision science, which is used to evaluate sensitivity to contrast. A participant might be presented with one target, and asked to indicate if the lines are oriented to the left or right. At the lowest level, Figure 4.1 (1), the intensity of the stimulus is so low that an accurate left/right decision is impossible, and the participant can only guess at which orientation was presented. As intensity increases, detection and identification are easier and for the highest intensity (shown in Figure 4.1 (7)) there would be an expectation of 100% accuracy for accurately reporting the orientation of the stimulus.

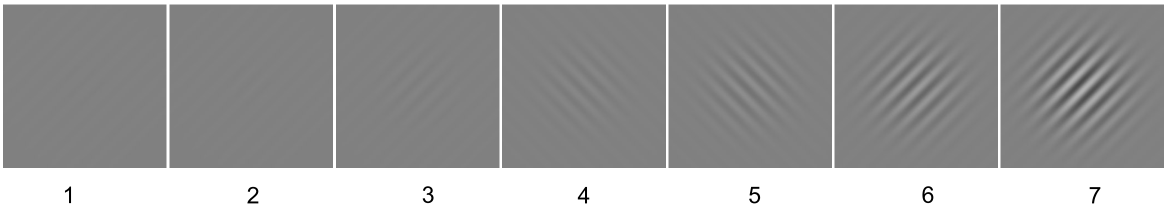


Figure. 2.1 Sample stimuli: participants would be required to make a left/right judgement on the orientation of the gratings. Gratings are varied in contrast (intensity) and adding more contrast to the gratings increases intensity of the stimulus, making a perceptual decision easier, and thus the probability of a correct response higher.

Responses are commonly collected using a forced-choice protocol. Forced-choice tasks require that a perceptual decision is always made (Kingdom & Prins, 2010). The most frequently used forced choice response protocol in pscyhophysics is the two alternative forced choice (2AFC). The number denotes the number of options the participant has with which to make their perceptual decision, thus in a 2AFC one response is provided with a choice of two options. In this way, 2AFC can refer to the number of response options for a trial, such as left/right, top/bottom, or first interval/second interval (Kingdom & Prins, 2010). After presenting the stimulus, such as those used in Figure 4.1, the participant makes a left/right response. After repeated trials and responses, the average response for each level can be calculated and plotted on a graph. In a 2AFC task of this type, the probability of a correct response at the lowest signal intensity is 50%; since there is not a clear signal, the expectation is for a random response. Therefore the expected range of responses in a 2AFC detection task is between 50% (chance) where stimulus intensity is at its lowest to 100% (always correct) at the highest stimulus intensity (Kingdom & Prins, 2010; Watson & Fitzhugh, 1990). In rare cases, such as for example the perception of depth or motion in the opposite direction to that which is predicted (Aoki et al., 2017; Asher & Hibbard, 2018; Doi et al., 2011; Read & Eagle, 2000; Tanabe et al., 2008) the proportion of correct responses might drop below 50%, indicating a reliable tendency to respond with the incorrect answer. In other cases, the expected range of responses of a particular type is not between 50% and 100%, but between 0% and 100%. For example, in an orientation sensitivity task, a grating may be gradually rotated from anti-clockwise to clockwise relative to vertical across stimulus levels, and the participant may be asked to judge which of these two possibilities had been presented. We would then expect all anti-clockwise responses (0% clockwise responses) at one end of the spectrum, and 100% clockwise responses at the other. For generality, we will refer to ‘positive’, rather than ‘correct’ responses to cover both of these cases (Kingdom & Prins, 2010). While 2AFC is used extensively in psychophysics, the number of alternative

Table 2.1 Comparison of forced choice methods (Wichmann & Hill, 2001a). Chance performance (also called the guess-rate; γ) is the probability of a correct response should an participant guess on a trial. The guess-rate is the inverse of the number of alternatives in the forced-choice. Thus, in a 2AFC the guess-rate $(\gamma) = \frac{1}{2}$ hence a probability of 0.5 or a 50% chance of choosing either response. The threshold is most commonly defined at the halfway point between chance performance and perfect accuracy (100% correct).

Alternatives	**Chance**	**Threshold**
1AFC	0	50%
2AFC	50.0%	75.0%
3AFC	33.3%	66.7%
4AFC	25.0%	62.5%
5AFC	20.0%	60.0%
6AFC	16.7%	58.3%
7AFC	14.3%	57.1%
8AFC	12.5%	56.3%

choices is (potentially) unlimited, however this will change the level at which a researcher defines their threshold (see Table 2.1).

Plotting the proportion of correct responses as a function of intensity (see Figure 2.3 A) across the range of responses provides an opportunity to establish the participant's ability to reliably perceive and categorise stimuli. Fitting a smooth curve to these data, called a psychometric function, maps the relationship between the continuum of stimulus intensities and the participants' responses. The psychometric function is determined by mathematically fitting the best sigmoid (S-shaped curve) through the data set of each participant (see Figure 2.3 B). The S-Shape reflects the increasing tendency to make positive responses as the level of the stimulus increases. This can for example be modelled as a cumulative normal distribution (see Figure 2.2 A and B). After fitting the psychometric function it becomes possible to estimate the threshold, or any other level of performance at any stimulus level (Kingdom & Prins, 2010).

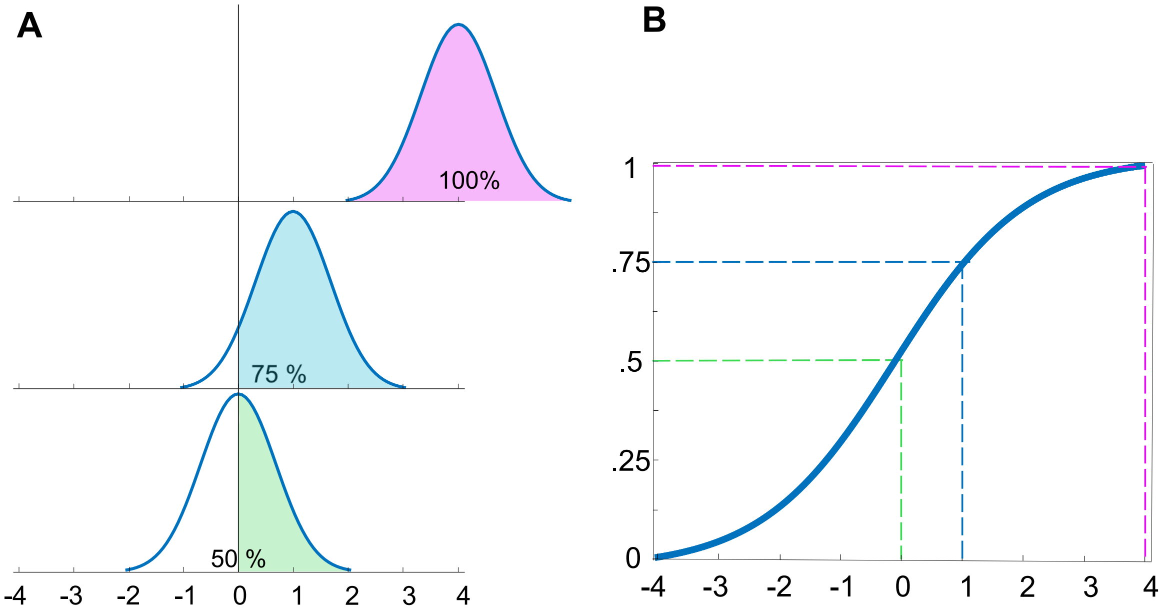


Figure. 2.2 (A) The normal distribution can be used to model the variability in responses to a given stimulus. (B) The cumulative distribution is the probability that any given variable will be equal to or less than any given value, and here defines the probability of a positive response as a function of the stimulus intensity.

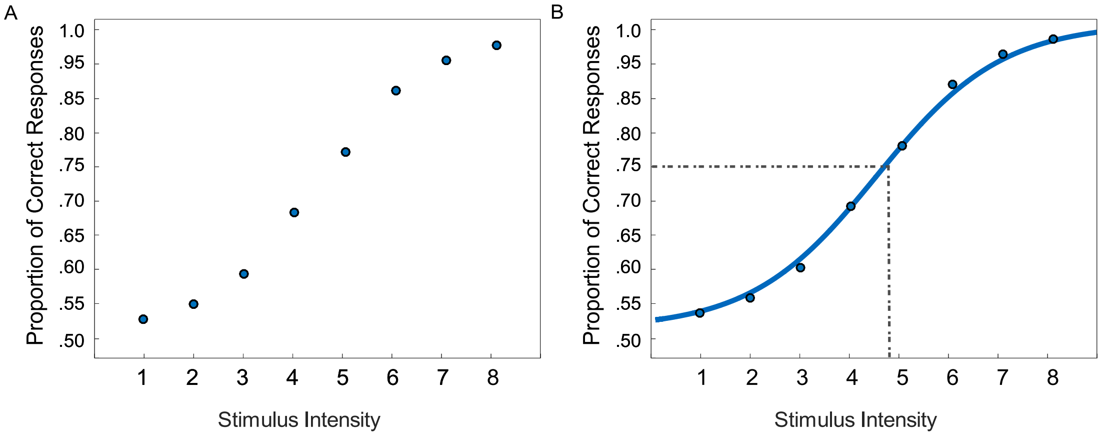


Figure. 2.3 Plotting the proportion of correct responses as a function of stimulus intensity shows that as intensity increases, so does the proportion of correct responses.

In a 2AFC task the intensity of the stimulus that is required to obtain an accuracy of 75% correct is often defined as the *threshold*. Thus, if a participant has a high sensitivity, they can reliably detect the target at lower stimulus intensities and therefore they will have a lower threshold (see Figure 2.4 (participant 1)). Conversely, a participant with low sensitivity requires high stimulus intensities to reliably detect the target, and they will have a higher threshold (see Figure 2.4 (participants 2-3)). The value of 75% is chosen for convenience since it sits halfway between chance-level and perfect performance.

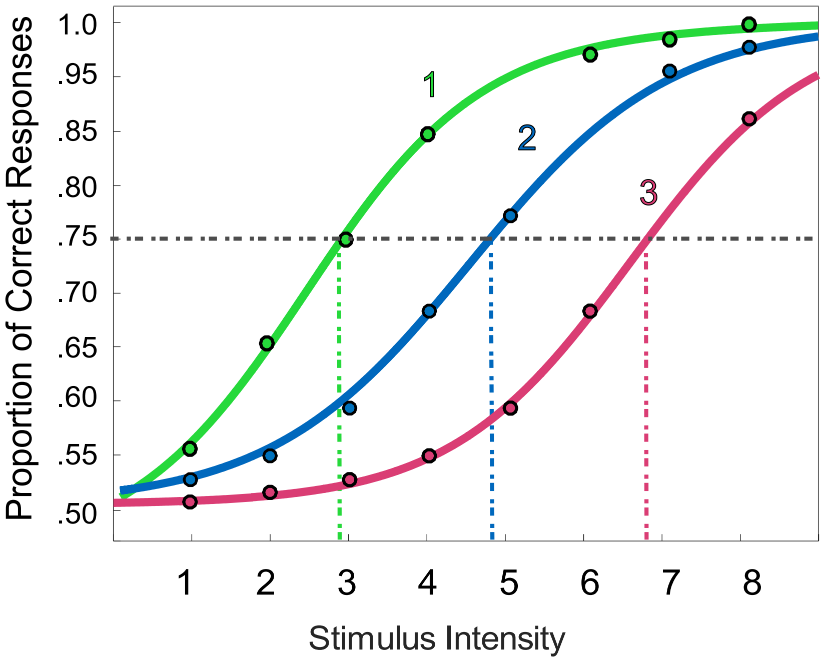


Figure. 2.4 Plotting the proportion of correct responses as a function of stimulus intensity shows that as intensity increases, so does the proportion of correct responses. The threshold is often defined (in a 2AFC task) as the point at which an participant obtains a correct response 75% of the time which will vary between participants. (1) A highly sensitive participant will require a less intense stimulus to achieve 75% accuracy, compared to an participant who has an average (2) or low (3) sensitivity.

2.3 Fitting the best Sigmoid

The psychometric function defines the increasing likelihood (or probability, p) of a positive response with increasing stimulus strength (x) (Gescheider, 1997; Kingdom & Prins, 2010; Treutwein & Strasburger, 1999) by generating a curve that best fits the participants' experimental data (Kingdom & Prins, 2010). In order to fit this curve effectively, it is important that data are collected in such a way that its shape is clearly defined. Stimuli that are vastly above or below the expected level of detection may provide an inaccurate measure of the threshold and invalidate the threshold estimation (Kingdom & Prins, 2010; Wichmann & Hill, 2001a,b). The first steps to ensuring an accurate psychometric fit are (i) a comprehensive range of levels of the stimulus intensity, (ii) a suitable spread of the intensities within this range, and (iii) enough repetitions at each level to provide an accurate estimate of the proportion of positive responses.

2.3.1 Range of Levels

In order to accurately fit a curve to experimental data a comprehensive range of levels of stimulus intensity are needed. Levels that are vastly above or below the expected level of detection may provide an inaccurate measure of the threshold. If the intensities sampled are all too easy or too difficult to detect or discriminate, then the psychometric fit cannot extrapolate an accurate estimate of any other value. Figure 2.5 show examples of a task that has been set at a level that is too easy (green curve) or too difficult (pink curve) to allow a 75% threshold to be estimated. Levels should be carefully selected in order to elicit a range of responses from chance (50% correct) to perfect (100% correct) (Kingdom & Prins, 2016).

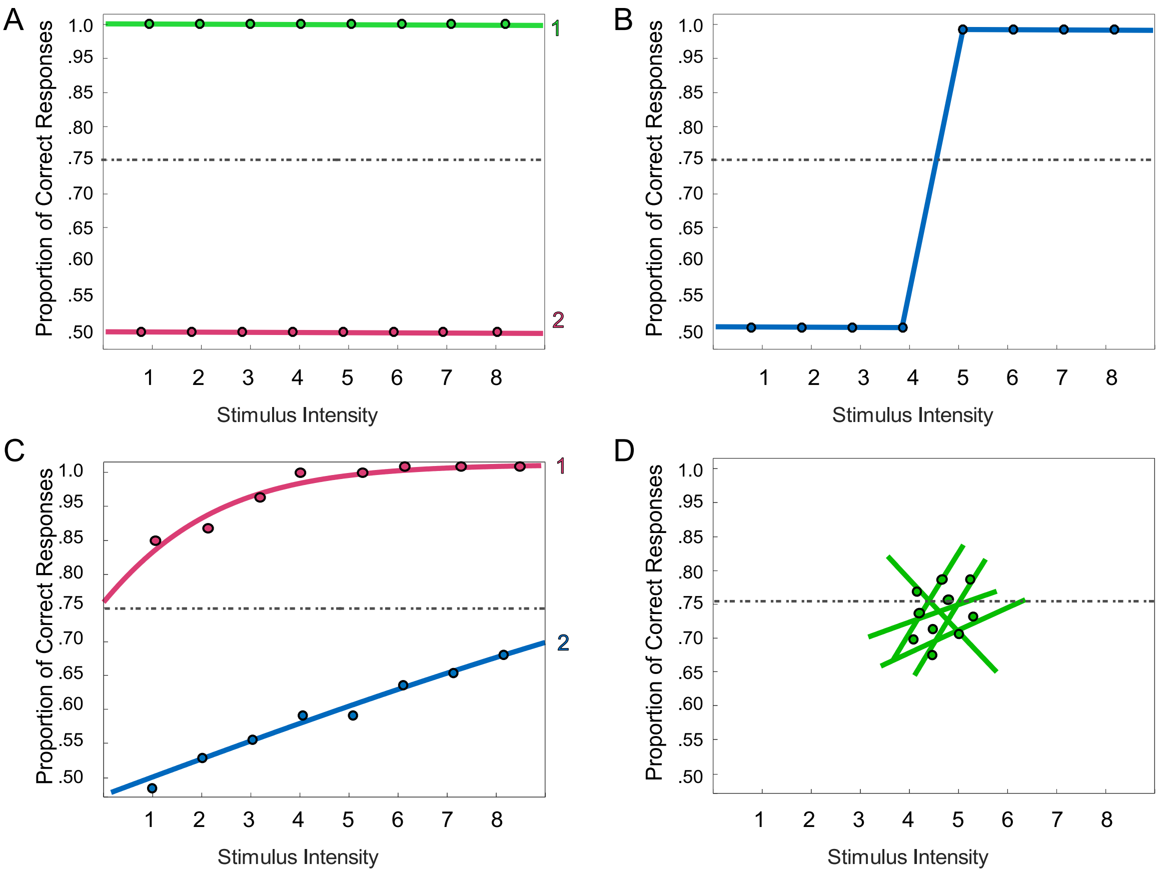


Figure. 2.5 Psychometric functions where the range or spread of levels do not provide sufficient data in order to reliably estimate a threshold it is necessary that the range of relevant stimulus levels is sampled appropriately. **(A)** The stimulus levels are either all too difficult (pink curve) or too easy (green curve) meaning that there is no variation in response with stimulus level, so the shape of the psychometric function cannot be estimated **(B)** Easy and difficult trials are included, but there is a sudden jump from chance to perfect performance, meaning that the threshold level cannot be determined precisely **(C)** There is variation in performance with stimulus level, but all the trials are either too easy (pink curve) or too difficult (blue curve) such that in neither case do the sampled levels span the 75% threshold, which cannot then be estimated without extrapolation. **(D)** All stimuli are presented at or around the 75% threshold, within a very narrow range. As a result there is no discernable relationship between stimulus level and performance, and it is not possible to estimate the slope of the psychometric function.

2.3.2 Spread of Levels

Even if the stimulus levels chosen capture the full range of performance from chance to perfect, they may still be insufficient to accurately constrain threshold measurement. An example is shown by the blue curve of Figure 2.5 (B). In this case, there is a sudden transition from chance to perfect performance, between the fourth and fifth levels presented. The 75% level will lie somewhere within this range, but the sampling of the stimulus intensity scale is not fine enough to allow the threshold to be estimated precisely.

2.3.3 Number of Repetitions

Finally, the psychometric function is extremely sensitive to errors that are not due to the intensity of the stimulus, but to natural participant variability from trial to trial. A good number of repetitions is needed to reduce the variability in the estimate of the proportion of positive responses at each level. The required precision of the estimate and the number of levels will dictate the number of trials required, however Kingdom and Prins (2016) suggest that in order to reliably estimate the threshold and the slope, approximately 400 repetitions is a plausible target.

2.4 Data Collection Methods

The two primary methods for presenting the stimulus and capturing participant response are the Method of Constant Stimuli (MOCS) and Adaptive Procedures (staircases).

2.4.1 Method of Constant Stimuli

MOCS involves repeatedly presenting a randomly selected stimulus from a predefined set of levels (Kingdom & Prins, 2010) and the likelihood of positive responses increases with the increase in the stimulus level. Plotting the proportion of positive responses as a function of level would resemble (in ideal situations) Figure 2.3 A and the experimenter could fit the psychometric function and estimate the threshold.

However, when employing MOCS as the experimental procedure, the three parameters ((i) Range of Levels, ii) Spread of levels, and iii) Number of Repetitions) need to be explicitly defined prior to each session of data collection, and there is no guarantee that a suitable set of stimuli will be chosen for each combination of participant and experimental condition.

Stimulus levels at the extremes provide very little useful data and waste trials (Watson & Fitzhugh, 1990). If a participant's threshold is not constrained within the range of defined stimulus intensities, (i.e. much higher / lower) they will either detect everything, or perform at chance (Figure 2.5(A)). It is therefore critical to use a sampling distribution that covers a comprehensive range of intensities (Gescheider, 1997). Furthermore, as illustrated in Figure 2.5 (B), the spread of levels may be too broad to constrain the threshold. Kingdom and Prins (2010) proposed that the critical requirement for accurately predefined signal intensity levels prior to experimental data collection can be alleviated with careful consideration and pilot testing. Watson and Fitzhugh (1990) argued that the use of MOCS as a method for threshold estimation is inefficient and overly time-consuming, and propose using adaptive procedures as an alternative (Watson & Fitzhugh, 1990).

2.4.2 Adaptive Procedures (staircases)

Adaptive procedures are another commonly used method of presenting stimuli, and do not require the prior categorisation of stimulus levels. In contrast, adaptive procedures use an algorithm to select the stimulus intensity for each trial. Depending on the type of staircase used, the algorithm analyses the responses from previous trials in order to converge on a predefined threshold (Kingdom & Prins, 2010; Shen, 2013). Fewer trials are required to home in on a participant's threshold, and thus the experiment is less time-consuming. There are many types of adaptive procedures that use a wide range of customisable algorithms (see Kingdom and Prins (2010) for a thorough overview). For the purpose of this thesis we will only consider an example of how staircases converge on a threshold. Staircases use rules to increase or decrease stimulus levels based on correct or incorrect responses to a defined number of previous trials. For example in a *1-up 2-down* detection staircase, stimulus intensity increases (gets easier) with every incorrect response, and decreases (gets more difficult) after 2 consecutive correct responses. The threshold (the signal intensity required by the participant to obtain the predefined performance level) on which the staircase converges will depend on the algorithm and the number of steps. In a one-step rule stimulus intensity will increase after an incorrect response, and decrease with a correct response and is referred to as simple up-down procedure (*1-up 1-down*) this staircase converges on a threshold of 50%. The *1-up 2-down* staircase converges on 70.7%, while a *1-up 3-down* staircase converges on a threshold of 79.4% (Levitt, 1971; Treutwein, 1995).

While adaptive procedures provide a more efficient way of obtaining a threshold, they do not provide reliable information about the slope of the psychometric function (Kingdom & Prins, 2010), because the data points are clustered around a single threshold value (see Figure 2.5 D). Wichmann and Hill (2001b) note that using a large number of trials that converge around a single estimate is inadequate to accurately define the slope of the curve and provide a reliable approximation for the model fit; Kaernbach (2001) demonstrated that bias in estimating the slope can result from the sequential dependence on the previous trial(s). In order to prevent this bias it is recommended to track more than one threshold level along the psychometric function, for example by using multiple interleaved staircases (Kaernbach, 2001; Shen, 2013), and to fit a psychometric function to the resulting data.

The choice of method will most likely depend on the individual researcher's preference. For a more detailed discussion on sampling for an optimal psychometric function, see García-Pérez and Alcalá-Quintana (2005). Ultimately, in order to produce an accurate estimation of the slope of the psychometric function the most reliable fit requires a comprehensive spread of responses across relevant stimulus levels.

2.5 The psychometric function

The psychometric function is an analytical tool that defines the increasing probability of a positive response (p) with increasing stimulus level (x) (Gescheider, 1997; Kingdom & Prins, 2010; Treutwein & Strasburger, 1999). The psychometric function is determined by mathematically fitting the best-fitting sigmoid through the data set of each participant in order to estimate the threshold (Kingdom & Prins, 2010).

In its simplest form, the psychometric function is defined by two parameters: the threshold (α (alpha)) and the slope (β (beta)). α determines the horizontal position of the curve along the x-axis (see **A** Figure 2.6 (α)). β determines the slope of the curve: how quickly positive responses increase with an increase in stimulus level (**B** on Figure 2.6 (β)). β provides a measure of the extent to which the participant's responses are affected by changes in the stimulus level.

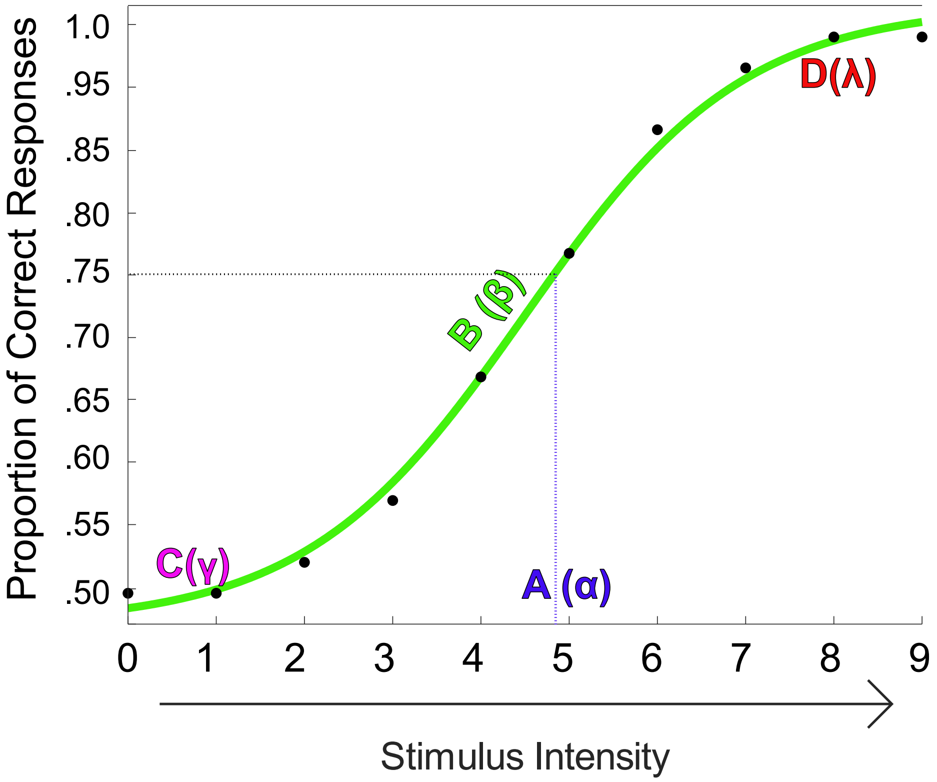


Figure. 2.6 The psychometric function: A(α) is the threshold, i.e. the intensity of the stimulus where performance is correct 75% of the time. B(β) is the slope, the spread of scores as a function of increasing stimulus intensity. C is the intercept and the lower asymptote, which illustrates performance at the lowest stimulus intensities. Finally, the upper asymptote located at D, describes performance at high stimulus levels where performance is expected to be near perfect.

2.5.1 Defining the psychometric function

In addition to i) the threshold (α) and ii) the slope (β) the equation requires two additional parameters to fully define the function in order to link the psychometric function to the experimental data: iii) the lower asymptote λ (lambda) and iv) the upper asymptote γ (gamma). (C and D in Figure 2.6 respectively).

These parameters relate to the guess-rate and lapse-rate respectively (Kingdom & Prins, 2010). The guess-rate is the probability of a correct response, should a participant guess on a trial. The guess-rate is the inverse of the number of alternatives in the forced-choice. In a 2AFC the guess-rate $(\gamma) = \frac{1}{2}$ is 0.5. The lapse-rate (λ) is the probability of a participant error, even when the stimulus intensity is well above their threshold. When the stimulus is at its highest intensity, performance is expected to be perfect (100%), it is common in forced-choice experiments for λ to be fixed at a probability of 0. However, this assumes that there will be no errors that are independent of the intensity of the stimulus, such as a participant sneezing or dozing. The fitted psychometric function is vulnerable to errors that do not originate from the stimulus intensity. Any incorrect responses (lapses) at the highest intensity can significantly bias the estimation of the curve fit (Kingdom & Prins, 2010; Wichmann & Hill, 2001b) and result in the curve asymptoting at a value less than 1. Once the lapse-rate has been determined, the upper asymptote is defined as $1 - \lambda$ (Kingdom & Prins, 2010). There are a number of curve functions that can be used to model the data, such as the logistic, Weibull, Gumbel and cumulative Gaussian (Kingdom & Prins, 2010; Wichmann & Hill, 2001b). These functions all reflect the typical sigmoidal shape of the data. However, the equations that define each function differ and thus the estimation of the curve will also differ slightly, depending on the type of function chosen (Kingdom & Prins, 2010). Toolboxes such as Palamedes (Kingdom & Prins, 2016; Prins & Kingdom, 2018) are available to fit psychometric functions, in order to fit the psychometric function, estimate the parameters of the fit and perform comparisons of models.

2.5.2 Building the generic psychometric function

Non-linear curve fitting can be used to fit the psychometric function ($\psi(x)$) to empirical data, and provide estimates of the four parameters (($\alpha, \beta, \gamma, \lambda$):

$$\psi(x) = \psi(x; \alpha, \beta, \gamma, \lambda)$$

where x is the stimulus level. The psychometric function requires all four parameters, including γ and λ, to be defined. These parameters are defined as the probability of a correct guess γ when a stimulus is unseen; either due to a lapse λ or due to the stimulus being below their threshold of detection, 1 - λ or not a lapse. When a stimulus is not seen due to participant error (λ) but guessed correctly (γ), the probability of a correct response is the product of the two parameters. Figure 2.7 illustrates the probability path for a decision, and relates the four parameters to the psychometric function. There are three scenarios for a correct response which form the sum of the probabilities for the guess- and lapse-rate and two only two that are dependent on the intensity of the stimulus, therefore:

$$\begin{aligned} \text{no lapse, correct, no guess:} \quad & p = (1 - \lambda) \quad + \\ \text{no lapse, correct guess:} \quad & p = (1 - \lambda) \cdot (\gamma) \\ & \sum p = \gamma + (1 - \lambda - \gamma) \end{aligned} \tag{2.1}$$

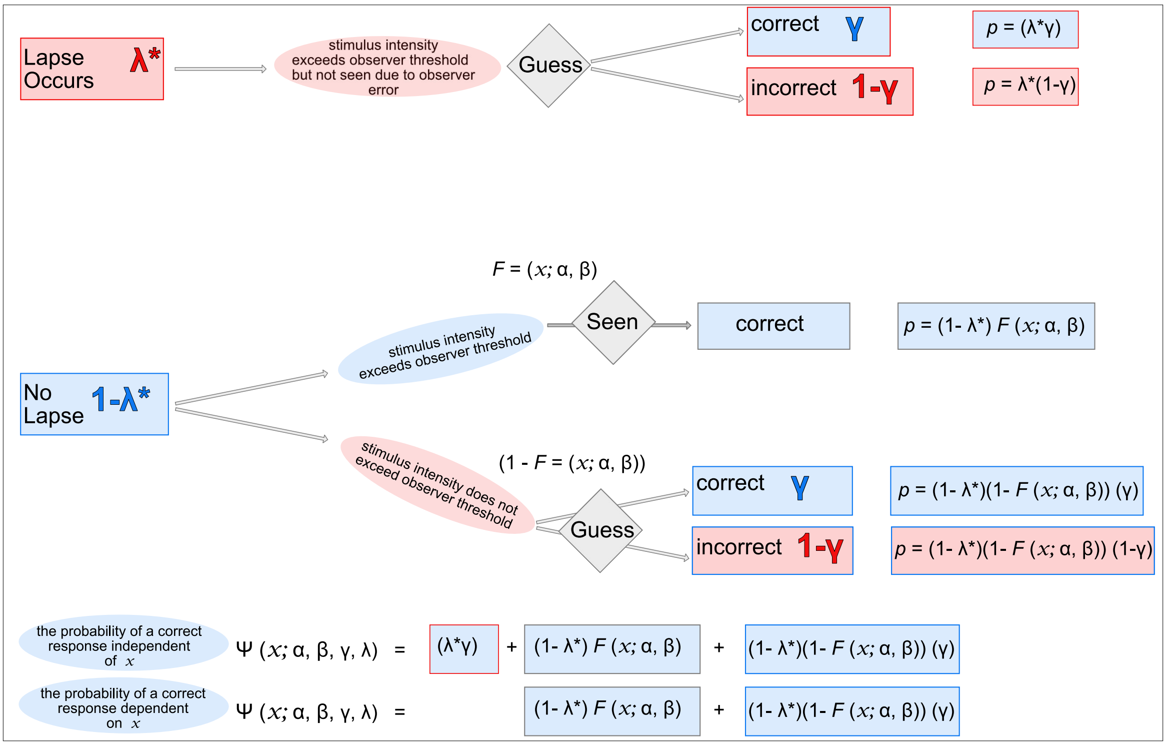


Figure. 2.7 Decision tree, adapted from Kingdom and Prins (2010, p.77) that illustrates the relationship between the four parameters of the psychometric function. Where β (beta) and α (alpha) are the slope and position of the curve respectively. The guess-rate γ (gamma) is the probability of a correct response should an participant guess on a trial and λ^*) (lambda) is the probability of a correct response after lapse. A lapse is defined as a stimulus-independent guess.

Finally all the elements are combined to produce a full psychometric function plotted alongside the raw data to evaluate the how accurately the psychometric function fits the data.

$$\psi(x;\alpha,\beta,\gamma,\lambda) = \gamma + (1-\gamma-\lambda)\cdot F(x;\alpha,\beta) \tag{2.2}$$

2.6 The psychometric function and individual differences

One goal of psychophysics is to understand the nature of the neural mechanisms underlying perception (Wilmer, 2008). The psychometric function is a powerful tool to estimate thresholds and characterise how each participant's responses vary with the stimuli presented, and is a commonly used method to establish differences in thresholds between control (normal) and experimental (abnormal) populations. This strategy has been used to investigate and understand the differences in neural processing for conditions such as migraine (Asher et al., 2018a), M.E. (Ahmed et al., 2018) amblyopia (Camilleri et al., 2014a) and macular degeneration (Plank et al., 2014). These comparisons are made at the level of the individual, and once the psychometric function has been fit for each individual participant the threshold values are accumulated and averaged to identify group level differences (Moscatelli et al., 2012). Since members of a species have the same physiology, the assumption is often made that this will produce homogeneous behavioural results that generalise to the broader population (Kosslyn et al., 2002; Nefs et al., 2010; Wichmann & Hill, 2001b).

However, even individuals within the normal population will differ, naturally, in their abilities to perceive and perform on perceptual tasks (Halpern et al., 1999). Where an individual's performance is not at the level required to provide an accurate estimate of the 75% threshold, an unreliable estimate may be provided, thus increasing the variability and uncertainty in the overall results. Data that do not provide a reliable estimate of threshold either are often described as non-normal, abnormal, atypical or poor. Several strategies are employed to ensure that extreme inter-individual variability does not distort the final results. Experimenters sometimes select their participants by 'pre-screening' them to ensure that they 'can do' the task (Heron & Lages, 2012). These screening processes may result in a sampling bias. In the area of stereo vision there is a high variability in individual performance, and within a typical sample deficits are common (Heron & Lages, 2012). Heron and Lages (2012) found that almost 50% of published studies between 2000-2008 did not report their screening processes and 91% did not report their selection criteria, potentially creating a wide-spread sampling bias in binocular vision research (Heron & Lages, 2012). Additional strategies used to normalise variability include training participants on the task before commencing the experiment. However, extreme performance on either end (good or bad), either before or after training, can provide valuable information about perception. There is accumulating evidence that individual variation in performance can provide robust insights into the underlying neural mechanisms that underpin perceptual behaviour (Baldassarre et al., 2012; Baldassarree et al., 2016; Hibbard et al., 2002; Kosslyn et al., 2002; Mollon et al., 2017; Nefs et al., 2010; Peterzell et al., 2018). Finally, if a threshold cannot be estimated in one condition, then it may be necessary to exclude that participant entirely from statistical analysis. This reduces the power of the sample, and runs the risk of biasing the conclusions.

Describing variability as undesirable or invalid implies these data are not worthy of analysis (Kosslyn et al., 2002). In order to reduce variability and minimise outliers, the problematic data are often adjusted, transformed or excluded from the analysis, some of which may go unreported in publication. Data manipulation techniques, such as transformations in order to achieve homogeneity, are also employed to normalise the variability within heterogeneous data (Bolker et al., 2009). The complexity in achieving the homogeneity required for this kind of analysis may provide another reason why smaller sample sizes are more common within psychophysical research.

2.6.1 Inter-individual differences

Mollon et al. (2017, p.4) argue that there are two important distinctions to be made regarding individual differences, firstly ‘differences in the *data*’ and secondly, ‘*true* individual differences’. Where differences in the data, are defined as those that arise from measurement error, random variability, external noise, systematic biases etc. In contrast, true individual differences, such as age and gender, are intrinsic differences that Mollon et al. (2017, p.5) defines, “as a hypothetical construct”. those that remain after all the measurement error has been excluded. The authors provide useful suggestions on how to design experiments to investigate these differences in more depth. However, for the purpose of this discussion we focus on the former. Individual variability in performance has been reported in colour vision (Bosten et al., 2014), stereoacuity (Schmitt et al., 2002), depth perception (Asher & Hibbard, 2018), motion in depth (Nefs et al., 2010), perceptual learning (Aberg & Herzog, 2009; Fahle, 2005; Jeter et al., 2009; Schmitt et al., 2002) and several other visual (Halpern et al., 1999) tasks. Where larger samples are used, natural inter-individual variability in performance can provide a novel insight into the mechanisms involved.

One example, from the field of depth perception, used individual differences to provide an insight into how the visual system perceives objects moving in depth (Nefs et al., 2010). Understanding how binocular vision contributes to the perception of motion in depth has been a debated topic (Harris et al., 2008; Nefs et al., 2010). While there is independent evidence for two theoretical mechanisms (based on changes in binocular disparity or binocular differences in velocity), these are often ambiguous and/or contradictory due to the difficulty of separating the two sources of information in experimental stimuli (reviewed by Harris et al. (2008)). However, utilising an approach that accounted for individual differences in a normal population, Nefs et al. (2010) demonstrated that both mechanisms were present. Furthermore, their results showed that depending on the type of visual cue used, some participants only obtained a reliable threshold for one cue, others were able to establish a reliable threshold for both, and finally some of the population did not establish a threshold for either. Importantly, Nefs et al. (2010) highlight that when using larger sample sizes there is an increased risk that some participants will not perform well, however this is representative of a normal population. Other studies have similarly used individual difference to understand the mechanisms of binocular depth perception (Hibbard et al., 2002; Peterzell et al., 2018).

Inter-individual differences have been reported in studies investigating learning and transfer (Aberg & Herzog, 2009; Fahle, 2005; Fahle & Henke-Fahle, 1996; Jeter et al., 2009; Schmitt et al., 2002) and while it is generally expected that performance would improve with increased training, there is evidence of an inverse relationship between the rate of learning and a participants' initial performance (Astle et al., 2013; Baldassarre et al., 2012; Baldassarree et al., 2016; Fahle, 1997; Fahle & Henke-Fahle, 1996). Furthermore, some training programmes work for some individuals, but not others. Schmitt et al. (2002) showed significant inter-individual variability in a normal population when training participants to improve their stereoacuity where some participants improved by a large factor but others did not improve at all.

2.6.2 Intra-individual differences

In addition to inter-individual differences, experiments in perceptual learning have shown that there can also be variation in intra-individual performance. Because repeated measures protocols assess participants at multiple time points, individual performance may also vary at each point (Fahle & Henke-Fahle, 1996; Schmitt et al., 2002). Fluctuating performance increases the likelihood of an unreliable fit between the psychometric function and the data. In some experimental protocols data are collected through repeated measures from each participant, over several time points, such as in learning experiments, or when a large amount of time is required. In these cases, the cost of missing data is high, and can cause significant problems when performing the analyses. If participants miss a testing session, their results may have to be excluded from the analysis or in some case the missing data is replaced by a group average (replacement by mean or median), further reducing any variation in the results (Allison, 2010; Lachaud & Renaud, 2011).

2.6.3 The complexity of heterogeneous data

Embracing rather than excluding the inter- and intra-individual variation within datasets can contribute to a far more powerful and representative understanding of the broad population (Agresti, 2013; Baldassarre et al., 2012; Baldassarree et al., 2016; Bolker et al., 2009; Halpern et al., 1999; Kosslyn et al., 2002; Mollon et al., 2017; Moscatelli et al., 2012; Nefs et al., 2010). The argument that heterogeneous data are in fact useful, interesting, normal and more likely to be generalisable to the wider-population is a convincing one. However, in order to account for all this variability, a more sophisticated method is needed to model the data. Mixed-effects models, while not commonly used within psychophysics, are widely used in other research disciplines such as psycholinguistics (Barr et al., 2013), ecology (Bolker et al., 2009; Thiele, 2012), clinical medicine (Casals et al., 2014; Vaina et al., 2014) and memory and language (Baayen et al., 2008; Barr et al., 2013). Mixed effects models are especially useful for the analysis of data that contain repeated measures methods. Recently a number of articles using these methods to model perceptual data at the population level have been published, including but not limited to (Amitay et al., 2006; Asher & Hibbard, 2018; Asher et al., 2018a,b; Coates & Chung, 2014; Kattner et al., 2017; Knoblauch & Maloney, 2012; Liu et al., 2018; Moscatelli et al., 2012). In the previous section we considered how to estimate a threshold by fitting a non-linear curve to participant data using a psychometric function, after which results are pooled, averaged and analysed using standard statistical techniques such as ANOVA, t-tests or non-parametric methods (Bolker et al., 2009). In the next section we introduce how the Generalised Linear Mixed-effects Model (GLMM) can be used to fit a sigmoid to each participant capturing the variability across individuals, while simultaneously performing the analysis on the response variable. Furthermore, even in the absence of a reliable estimate of threshold these models produce an intercept and slope and thus provide valuable insight into the level of performance on a given task, and how this is affected by manipulations of the stimulus. Beginning with a general linear model, we will evolve the models step by step to illustrate how they can be used to analyse clustered categorical data.

2.7 Evolving the Models

2.7.1 Linear Models

Linear models (LM) such as simple or multiple regression models fit a straight line to two (or more) continuous variables in order to estimate the relationship between them. The assumption is that both the responses and the error are normally distributed, and that the response increases or decreases as a linear function of the predictor(s). In a simple regression, a response Y is modelled as **a weighted linear function** of the explanatory/independent variable(s) X (Knoblauch & Maloney, 2012).

The equations for the LM are as follows:

$$\begin{aligned} \text{single predictor} \quad & Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i && (2.3) \\ \text{multiple predictors} \quad & Y = \beta_0 + (\beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \dots) + \varepsilon_i && (2.4) \end{aligned}$$

where Y_i is response variable for participant i, β_0 is the intercept (i.e. the Y value when X = 0), X_{i1} is the first predictor (explanatory variable) for participant i, X_{i2} is the second explanatory variable for participant i, and X_{i3} is the third explanatory variable for participant i. The three explanatory variables here define a line in multi-dimensional space (in this case with four dimensions), whose direction is defined by the β parameters (β_1, β_2, β_3). Finally ε_i is the error for participant i. The regression model estimates a single line that best fits the relationship between each of the predictors and the response variable using ordinary least squares to estimate the fit parameters. Linear model equations can be simplified by writing them in matrix form,

$$\mathbf{Y} = \mathbf{X}\beta + \varepsilon \tag{2.5}$$

where **X** is a matrix that contains all of the values for X_1 to X_3, β is an unknown vector to be estimated and **Y** is the response variable to be predicted. An assumption of the general linear method is that continuous variables increase or decrease indefinitely, with normally distributed variation. This makes them unsuitable for modelling psychophysical data, in which the proportion of positive responses is restricted to lie between 0 and 1, and variation is expected to follow a binomial distribution. An extension of the linear model that can accommodate psychophysical data is the Generalized Linear Model (GLM).

2.7.2 Generalized Linear Models

The GLM extends linear models by allowing more flexibility in the distribution of the response variable, such as in situations where responses do not adhere to the assumptions required for a linear regression (Agresti, 2013; Knoblauch & Maloney, 2012). These models can be used in psychophysical experiments to analyse the responses of individual participants (Moscatelli et al., 2012). In the case of data from a 2AFC psychophysical experiment, a *link function* is needed to link the probability of positive or negative responses to the values of the underlying predictors. Along with the link function, the GLM defines the linear predictor of the independent variable(s) (for example; X_1, X_2, X_3), and the probability distribution of the responses Y (for example Poisson, binomial, exponential, logistic).

Defining the Link Function

The link function $g(\cdot)$ relates the non-normal response variable Y to the linear predictor η (Agresti, 2013):

$$\begin{aligned} \text{the linear predictor:} \quad & \eta = -\infty \rightarrow +\infty \\ \text{the binomial response:} \quad & y = 0 \rightarrow 1 \\ \text{generic link function:} \quad & \eta = g(Expected(y)) \text{ or } g(E(y)) \end{aligned} \tag{2.6}$$

Many types of link function can be used, depending on the outcome variables. The **logit** and **probit** link functions are two commonly used functions when using GLM for psychophysical data. The logit function is the natural log of the odds for the response variable (in a 2AFC task with proportions varying between 0 or 1):

$$\eta = \log\left(\frac{p}{1-p}\right) \tag{2.7}$$

where η is the expected value from the function of the linear predictor and p is the probability of a positive response (1). The probit link function is the inverse of the cumulative distribution function of the standard normal distribution:

$$\eta = \phi^{-}1(p) \tag{2.8}$$

where η is the function of the linear predictor and p is the probability of a positive response, and ϕ^{-1} is the inverse of the cumulative normal distribution.

Building the GLM

The response Y is modelled as a *weighted non-linear function* (defined by the *link-function*) of the independent variable(s) X (Knoblauch & Maloney, 2012). The assumptions for the GLM are that responses Y_i are independent cases that are independently distributed and, while they are not required to be normally distributed, the distribution is expected to fall within the exponential family of distributions (Knoblauch & Maloney, 2012). The same assumption applies to the error term:

$$\eta = g(\mathbf{X}\beta) + \varepsilon \tag{2.9}$$

In this way, the non-linear relationship between the predictors and the response variable is included into the model.

Both of the LM and the GLM thus far have been defined as fixed-effects models. The fixed-effects parameters estimate the effect of the explanatory variables on the response variable, are usually the primary focus for an experimenter and are the same for every participant (Moscatelli et al., 2012). Both of these models contain a single source of variability, the random residual variation that remains after the expected variation from explanatory variables has been accounted for. Analysis of data for populations of participants using these models require two stages, firstly to fit a function to individual responses and establish their thresholds, and secondly to pool all the data in order to conduct the inferential analysis on the resulting thresholds (for example an ANOVA or t-test). However, Moscatelli et al. (2012) argue that fitting the psychometric function to all the data, from all participants, and simultaneously conducting the inferential analysis, would be a more efficient method of analysis. This approach requires a generalised linear mixed-effects model.

2.7.3 Generalized Linear Mixed-Effects Models

Fitting a population level psychometric function needs to take into account the variability between participants, in addition to the variability within an individual participant. The *Generalized Linear Mixed-(Effects) Model* (GLMM) is a more advanced regression model which contains both fixed and random effects. Estimating the variability stemming from both of these effects provides a more flexible psychometric fit (Agresti, 2013; Moscatelli et al., 2012). The GLMM is an ideal tool for the analysis of psychophysical data, as it is able to model categorical responses from repeated-measures, forced choice designs simultaneously at the population level (Moscatelli et al., 2012), and is capable of taking into account the random variability that occurs between individuals.

Fixed-effects as previously covered include the core aspects of the experimental design such as stimulus intensity and condition (Agresti, 2013), where the variable contains all the levels of interest. In contrast, **Random-effects** are variables that only contain a sample of all the possible levels of interest. In psychophysics this is usually the participant.

Visualising the GLMM

It is unusual to think of psychophysical data as hierarchical, however in repeated measures designs, data will be clustered due to multiple data points being contributed by each participant. The fixed component estimates the effect of interest, and the random component

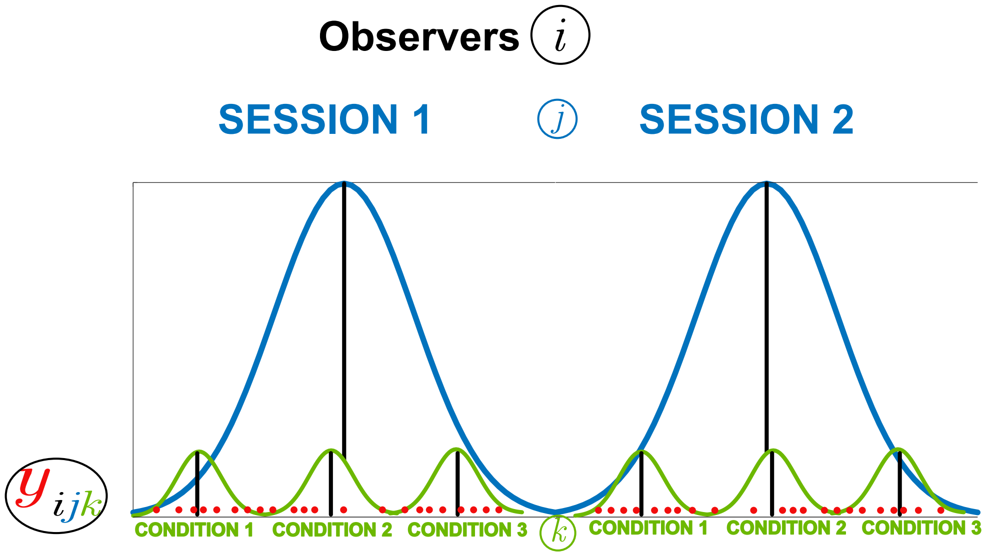


Figure. 2.8 Illustrative hierarchy of a data-set for a single participant out of n participants, adapted from Huber (2014). The individual responses y (red dots) deviate from its condition mean (y) and provide the within subjects variability given by the term ε_{ijk}. Hierarchically this defines the error for n observations by a single participant, within two sessions, and three conditions. The mean of each condition differs from the session j mean by Mean_{ij}. Finally the mean for each session differs within participant, Mean_{i}.

estimates the variation between the clusters of interest (in this case, between participants). Figure 2.8 shows a hypothetical scenario where Session and Condition are the independent variables (fixed effects), and the variability between the participants within these clusters make up the random effects. The fixed-effects model as defined previously in, Equation 2.9 ($\eta = \mathbf{X}\beta + \varepsilon$) can now be evolved to include the random-effects.

$$\text{mixed-effects model} \qquad \eta = \mathbf{X}\beta + \mathbf{Z}b + \varepsilon \tag{2.10}$$

where η is the linear response, $\mathbf{X}$ (as before) is the matrix of fixed-effect explanatory variables, β (as before) is a vector of fixed-effects coefficients and error $\mathbf{Z}$ is the matrix of the random-effects, b is the vector of random-effect coefficients and error. Finally, ε is the residual error, and the residual variation that is unaccounted for by the model.

Building the Error Term

The variability for the model is estimated at each level of the hierarchy. The fixed components (explanatory variables) model the variability for Session (j) and Condition (k). The random component in this example will be the between-subjects responses from each of the participants (i).

$$\begin{aligned} \text{within subjects variability:} \quad & \varepsilon_{ijk} \sim N(0, {\sigma_\varepsilon}^2) \\ \text{between subjects variability:} \quad & U_i \sim N(0, {\sigma_u}^2) \\ \text{Error term:} \quad & V_{ijk} \sim U_i + \varepsilon_{ijk} \end{aligned} \tag{2.11}$$

where ε is the error for the fixed effects (Session and Condition) and U is the error for the random effect (participants). $N(0, \sigma^2)$ indicates that the error is normally distributed with a mean of 0 (zero) and a variance of σ^2. The GLMM calculates the residual variability (that is not accounted for by the fixed effects) for the between subjects means (Moscatelli et al., 2012). This allows the estimation to be generalised to the whole population.

Building the Fixed and Random Effects

The generalized linear mixed effects model for a psychophysical task is created by combining the fixed and random effects:

$$\begin{aligned} \text{fixed effect: session} \quad & = \beta_0 + \beta_1 X_{ijk} + \varepsilon_{ijk} \\ \text{fixed effect: condition} \quad & = \beta_0 + \beta_2 X_{ijk} + \varepsilon_{ijk} \\ \text{fixed effects:} \quad & = \beta_0 + (\beta_1 X_{ijk} + \beta_2 X_{ijk}) + \varepsilon_{ijk} \\ \text{random effect: participants} \quad & = \beta_0 + b_2 Z_i + U_i \end{aligned} \tag{2.12}$$

Combining the elements to build the model

$$
\begin{aligned}
& Y_{ijk} = \beta_0 + (\beta_1 X_{ijk} + \beta_2 X_{ijk}) + (b_2 Z_i) + V_{ijk} \\
\text{in matrix form:} \quad & \mathbf{Y} = \beta\mathbf{X} + b\mathbf{Z} + V \\
\text{modelled via the link function as:} \quad & \eta = g\left(\beta\mathbf{X} + b\mathbf{Z} + V\right)
\end{aligned}
\tag{2.13}
$$

where η is the linear predictor, **X** (as before) is the matrix of fixed-effect explanatory variables. β (as before) is a vector of fixed-effects coefficients **Z** is the matrix of the random-effects, b is the vector of random-effect coefficients (defined in equation 2.12), and $\boldsymbol{V}$ is the error term (defined in equation 2.11).

2.7.4 Random effects: intercepts and slopes

The optimal model fit is the one that provides the best compromise between the goodness of fit to the observed data and the number of free parameters (Agresti, 2007). Allowing the intercepts and/or slopes to vary gives the model more freedom to fit an estimation to the observed data. In clustered categorical data the intercept is defined by the reference group (or variable). The parameters for each predictor in this case determine how the predicted values vary across the different categorical values of the predictor. For a linear predictor, the intercept is the predicted value of the dependent variable when the predictor has a value of zero, and the slope defines how a change in the fixed-effect (e.g. increasing stimulus intensity) produces an increase in the response (e.g. more correct responses). Models that include random slopes increase the models flexibility in fitting the data, by incorporating an estimate ofthe variability in each parameter to the maximum likelihood calculations, however they need more data in order to estimate accurately and robustly (Harrison et al., 2005).

2.7.5 Selecting the best model

Regrettably, a single model does not provide a simple solution, and it is important to assess the best combination of fixed and random effects when choosing a model for the data. Where the simplest model (containing only fixed-effects) provides a poor estimate of the observed data, including random-effects may provide a better estimate. The best-fit can be determined by testing random and fixed effects and comparing the model outputs (Moscatelli et al., 2012). However, introducing more random components increases the complexity of the model and the likelihood of overfitting. Overfitting occurs when the model is more complicated than the observed data support. The most common test for model selection is the Akaike information criterion (AIC). The change in AIC (or ΔAIC) determines the relative strength of a model (in comparison to another model). The AIC penalises more complex models, and this penalty increases as the number of parameters increases (see Agresti (2007, p.141) for further guidance).

2.7.6 Disadvantages of the GLMM

While GLMMs provide flexibility when modelling varied and complicated data sets, the complexity in selecting the parameters of the model, fitting and interpreting the data are not straightforward (Harrison et al., 2018). This opens mixed effects models to misuse (Bolker et al., 2009). Therefore, it is important when applying GLMMs to refer to best-practice guidelines (Agresti, 2007; Barr et al., 2013; Bolker et al., 2009; Harrison et al., 2005).

2.8 Non-linear Mixed-Effects Models

In some conditions, performance may never reach the theoretical upper asymptote of perfect performance, and the function will asymptote at a proportion of correct responses at a level less than 1. This may result in a poor psychometric fit of the observer response data using the GLMM, if an asymptote of 1 is assumed. To accommodate this, a nonlinear generalised mixed effects model (NLME) can be used. This model includes an additional free parameter in order to model variability in the asymptotic performance at high signal levels. This change in asymptote has been shown to be an important characteristic of performance perceptual learning studies: participants' responses may asymptote at a low level of performance early in training, with the asymptotic value increasing as a result of learning (Donovan & Carrasco, 2015; Herrmann et al., 2010; Ling & Carrasco, 2006).

The nonlinear regression provides three measures to assess a change in performance over time. Firstly, like the GLMM a leftward shift in the curve indicates an improvement in threshold (see Figure 4.3(a)). An increase in slope indicates an increase in the rate at which performance increases with increase signal level. Finally a change in the asymptote indicates a significant change to the performance at the highest level of stimulus intensity These changes are independent aspects of the psychometric function fit, and may not necessarily be congruent. For example, it is possible to obtain an increase in one measure and a decrease (or no change) in another.

A nonlinear mixed effects regression model can be performed with the the *nlmefit* function in Matlab, in this case described above using the following model:

$$p = 0.5 + \frac{(A + A_d)}{1 + e^{-(K+K_dS)(C-(C_0+C_{0d}S))}} \tag{2.14}$$

where p is the proportion of correct responses, A determines the asymptotic level of performance, K defines the slope and C_0 defines the threshold. d is the change in performance, where A_d, K_d and C_{0d} determine the change in asymptote, slope and threshold respectively. C is the coherence level, and S is a dummy variable, taking on values of 0 or 1 for the sessions or the variables of comparison.

Since this model does not provide an easily accessible output of the significance of the comparisons, it may be necessary to use bootstrap methods to obtain confidence limits.

Chapter 3

Contrast sensitivity and lateral inhibition in Migraine-with-Aura

3.1 Introduction

Migraine is a debilitating neurological disorder affecting approximately 14% of the population (Burch et al., 2015; Stovner & Andree, 2010), and attacks are often characterised by an aversion to light and sound (Headache Classification Committee of the International Headache Society, 2013). Some individuals with migraine experience sensory disturbances preceding their attack, known as migraine aura, which are most commonly experienced in the visual modality (Queiroz et al., 1997; Shepherd, 2000). While the pathophysiology of migraine is still not fully understood, these differences in visual processing suggest that the visual pathways may play a key role. Consequently, perceptual tasks are often used to explore the behavioural differences in brain functioning (Battista et al., 2011; McKendrick et al., 2006). Sensory processing differences between migraine and control groups have been found for a multitude of psychophysical tasks, including sensitivity to temporal and spatial contrast (Benedek et al., 2002; McKendrick & Sampson, 2009), colour and orientation (Palmer et al., 2000; Shepherd, 2000), global motion and form coherence thresholds (Antal et al., 2005; Ditchfield et al., 2006; McKendrick et al., 2001) and adaptation to prolonged stimulation (McKendrick et al., 2006; Shepherd, 2001, 2006). These effects on sensory processing are not restricted to the time of the attack itself, and differences in performance are also shown interictally (in between attacks) (McKendrick et al., 2017; O'Hare & Hibbard, 2016).

Migraine has been associated with a state of cortical hyper-responsiveness, the idea that the brain responds more strongly to sensory stimuli in migraine (Aurora & Wilkinson, 2007; Coppola et al., 2007; McColl & Wilkinson, 2000). Possible explanations, to account for the neurophysiological abnormalities found in migraine, include reduced inhibitory control between neurons (Palmer et al., 2000), an inability to ignore noise (O'Hare & Hibbard, 2016; Tibber et al., 2014) or preactivation from the brainstem and/or thalamus (Coppola et al., 2013). Many aspects of low-level visual processing rely on excitatory and inhibitory lateral interactions between neurons, for example visual search, border segmentation (Zhaoping, 2014), and the integration and segmentation of local features into globally coherent forms (Field et al., 1993). The current study assesses the extent to which these lateral interactions, both excitatory and inhibitory, might differ in their strength and spatial extent in migraine-with-aura.

3.1.1 Altered Lateral Interactions in Migraine

Physiological evidence supports the idea of impaired lateral interactions in migraine. Bridge et al. (2015) found reduced levels of the inhibitory neurotransmitter gamma-aminobutyric acid (GABA) in a group with migraine-with-aura and visual sensitivity and argued that the reduced GABA levels disrupt the coupling between excitatory and inhibitory activity in visual areas. Consistent with these reduced GABA levels, there is also psychophysical evidence that lateral interactions between neural responses may be different in migraine, although the pattern of results across behavioural tasks is inconsistent (O'Hare & Hibbard, 2016). Coppola et al. (2013) recorded visual evoked potentials (VEPs) using electroencephalogram (EEG) electrodes placed on the back of the head, while observers viewed radial windmill dartboard or partial windmill visual patterns. They found that for migraine observers, VEP amplitudes were reduced during the intericatal period compared to controls but were normalised during the ictal period (during an attack). This provides evidence for an imbalance between excitation and inhibition for individuals with migraine during the interictal period (Coppola et al., 2013). Furthermore, sensitivity to orientation differences is reduced in migraine (Tibber et al., 2006). This task depends on lateral inhibition as a mechanism for sharpening orientation tuning (Seriès et al., 2004), and orientation sensitivity is known to correlate positively with GABA concentration in the visual cortex (Edden et al., 2009). Other studies have reported no difference in inhibitory interactions in migraine (McColl & Wilkinson, 2000). Using contrast detection against a high-contrast background to assess the effects of contrast gain control, Battista et al. (2011) compared inhibitory interactions between migraine and control groups using the Chubb illusion (Chubb et al., 1989). This task involves judging the contrast of a central texture of varying contrast within an annular surround. When the surround is higher in contrast, the centre appears reduced in contrast. Surround suppression relies on lateral interactions between neurons (Chubb et al., 1989). Increased surround suppression was found for moving stimuli in the migraine groups compared to the controls, but no difference was found for static targets.

3.1.2 Masking in Migraine

Visual masking involves introducing noise to a stimulus so as to make the target harder to detect. Masks can be presented before (forward masking) or after (backward masking) the target. If spatial inhibitory interactions in migraine groups are reduced, then the degree of masking would also be expected to be reduced (Breitmeyer & Ogmen, 2000). However, some evidence indicates that backward masking is no different in migraine groups in comparison with control groups (Palmer et al., 2000; Shepherd et al., 2011). Wagner et al. (2013) found increased masking in migraine-with-aura and attributed this to increased levels of multiplicative noise.

3.1.3 Global Processing in Migraine

Migraine groups also show a deficit in global processing tasks such as global form (McKendrick et al., 2006) and global motion detection (Tibber et al., 2014). Global form perception relies on the effects of an association field, a local network of inhibitory and excitatory interactions that group together individual elements in order to detect extended contours (Field et al., 1993). The association field is important in allowing for the detection of globally defined shape (Field et al., 1993) at higher stages of visual processing. The impoverished perception of global shape and form in migraine are consistent with an altered pattern of inhibitory and/or excitatory lateral interactions.

3.1.4 Technique for assessing Lateral Interactions

Physiological and psychophysical evidence suggest that lateral interactions may differ in migraine, and these differences are likely to be related to the reduced ability to detect global form. The aim of the current study was to use a detailed psychophysical approach to directly assess the strength and spatial extent of these interactions in migraine-with-aura. We used the technique pioneered by Polat and Sagi (1993) to do this. Observers were required to detect a central Gabor at varying levels of contrast, presented either alone or in the presence of high-contrast, colinear flankers of the same orientation and spatial frequency as, and spaced between 2 and 10 wavelengths from, the central target. Polat and Sagi found reduced contrast sensitivity when the flankers were very close to the target, consistent with visual masking effects, but improved sensitivity for flankers at a distance of up to 10 wavelengths, indicating the presence of local excitatory interactions. This design will therefore allow us to determine whether inhibitory and excitatory interactions differ in their magnitude or spatial extent in migraine-with-aura.

3.2 Materials and Methods

3.2.1 Participants

A total of 76 observers were tested. The categorisation of observers into groups was undertaken using the criteria of the Headache Classification Committee of the International Headache Society (2013). All observers completed the experiment regardless of group. However, because differences in masking have been found to be greater in migraine-with-aura than in those without aura (Wagner et al., 2013, 2010), only the with-aura group was investigated. Inclusion as a control observer required no symptoms of severe headache, migraine or aura. Migraine observers were tested interictally and were required to be free from migraine for 3 days before and after of the day of testing. No observers used prophylactic medication for migraine. The data for two migraine observers were excluded as a result of experiencing an attack within 3 days of their testing day. The duration of testing was for 50 min on one day only. After the classification process, there were 31 controls (19 females, mean age of 22.8) and 24 with migraine-with-aura (22 females, mean age of 26.6; see Table 3.1); 18 observers were excluded after being assessed as either migraine without aura, non-headache-free controls or migraine-with-aura not meeting inclusion criteria. All experiments were conducted in accordance with the World Medical Association Declaration of Helsinki (2013) and were approved by the University of Essex ethics committee (Application No. VR1403). All observers gave written, informed consent and received payment for their participation.

Table 3.1 Migraine observers' reports of clinical features.

Observer	**Sex**	**Age**	**Age Started**	**Prior Attack**	**Frequency per Month**
OB8	F	23	16	1 month	1–3
OB12	F	27	10	2 weeks	1–3
OB14	F	60	19	3 months	<1
OB16	F	52	14	5 days	1–3
OB17	F	22	17	5 days	1–3
OB20	M	30	12	4 months	5 or more
OB21	F	41	24	9 days	5 or more
OB22	F	21	10	4 days	<1
OB24	F	36	23	7 days	1–3
OB25	F	29	3	20 days	1–3
OB28	F	30	13	3 months	<1
OB29	F	55	40	6 days	1–3
OB30	F	25	9	8 days	1–3
OB31	F	23	19	1 month	1–3
OB33	M	20	12	7 days	1–3
OB48	F	18	NDI	1 month	5 or more
OB50	F	21	NDI	7 days	<1
OB63	F	18	13	1 month	1–3
OB64	F	19	17	7 days	5 or more
OB66	F	19	16	A few months	1–3
OB71	F	20	11	> 3 days	<1
OB73	F	18	15	6 days	5 or more
OB75	F	19	16	3 months	<1
OB76	F	18	NDI	> 3 days	NDI
Excluded	F	19	16	1 day	1–3
Excluded	F	18	14	2 days	1–3

NDI—non-disclosed information.

3.2.2 Apparatus

Stimuli were presented using on a Sony Trinitron 2100 monitor with a screen resolution of 1280 × 1024 pixels and a vertical refresh rate of 100 Hz. One pixel subtended 0.7 arc min. A Datapixx CRT Driver (Vpixx Technologies, Saint-Bruno, QC, Canada) was used to achieve 16-bit control of contrast levels. Stimuli were generated and presented using MATLAB and the Psychophysics Toolbox extensions (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997). Responses were made via the left and right arrow keys on a standard keyboard.

3.2.3 Stimuli

Stimuli were presented on a mid-grey background. The target stimuli were centrally presented, vertically oriented Gabor patches, with a spatial frequency of 4 cycles per degree. The contrast of the target was manipulated: there were seven contrast levels (0.0075%, 0.015%, 0.0225%, 0.03%, 0.045%, 0.06% and 0.09%; Michelson Contrast). On a logarithmic scale the step difference(s) between levels was 0.0075 for the first three increments, 0.0150 for the second two, and 0.03 for the final increment. This provided a combination of relatively fine tuning for the lower contrast levels, and a broad coverage of contrasts. The flankers consisted of two colinear high-contrast, high-spatial-frequency Gabors with a spatial frequency of 4 cycles per degree. Flankers were positioned above and below the target at one of six distances (1, 2, 3, 4, 6 or 12 wavelengths) from the target. A baseline condition was also included, in which the target was presented alone (without flankers). See Figure 3.1 for stimuli.

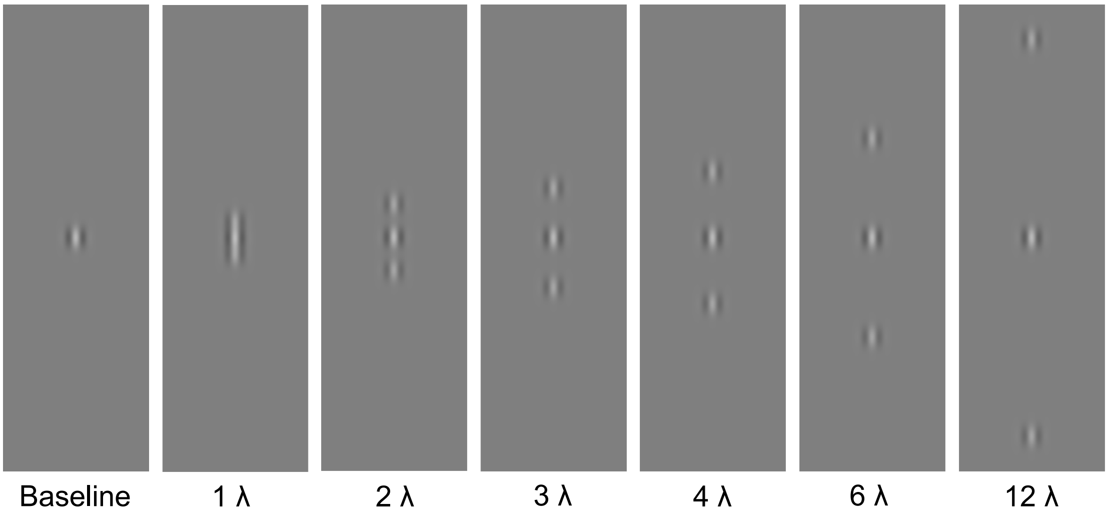


Figure. 3.1 Experimental stimuli. Central Gabor target presented alone, and in the presence of high-contrast, high-spatial-frequency flanking Gabors at one of six distances (1λ, 2λ, 3λ, 4λ, 5λ, 6λ or 12λ).

3.2.4 Procedure

Observers were positioned at a viewing distance of 1.5 m from the display. The task consisted of a two-alternative-forced-choice (2AFC) procedure to identify which of two intervals contained the target. The flankers (if present) appeared in both intervals (Figure 3.2). Stimulus presentation began with a central fixation cross displayed for a duration of between 670 and 800 ms, randomly selected for each trial. The first interval was presented for 90 ms, followed by a central fixation cross for a randomly selected duration of between 270 and 400 ms, followed by the second interval also presented for 90 ms. During the stimulus presentation, additional crosses were presented either side of the central target to alert the observer that a presentation was in progress and to aid with identifying where the target would appear. These crosses were located at 70 arc min to the left and right of the central target location. There were 20 repetitions of each contrast level, resulting in a total of 140

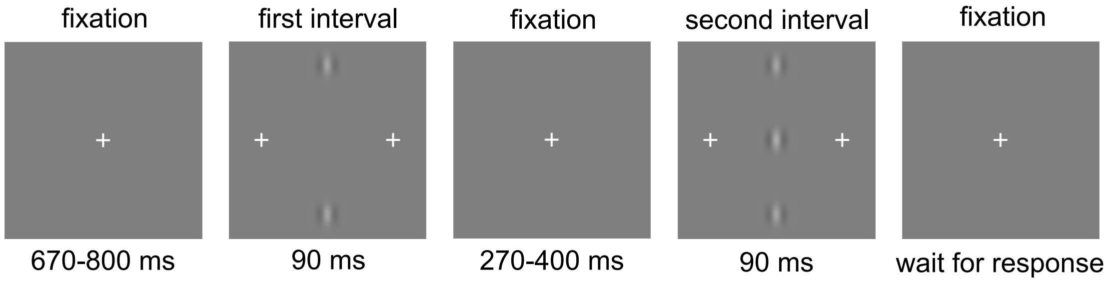


Figure. 3.2 Observers were required to select which of two presentations contained the central Gabor target. In this example, the target is presented in the second interval.

trials for each flanker distance. Each distance was presented in a separate block, resulting in seven blocks, one for each of the six flanker distances and one for the baseline (no flanker) condition. Contrast levels were randomised within each block. The block order was also randomised for each observer.

3.3 Results

3.3.1 Analysis

The responses from each observer were converted into the probability of being correct out of the total number of trials for each condition. A generalised linear mixed effects model was used to analyse the data, using the *fitglme* function in MATLAB. We first compared the contrast sensitivity for the Gabor target in the absence of flankers, using a model with contrast as a fixed covariate, group as a categorical factor, and observer as a random factor, with random intercepts and contrast slopes. A probit link function was used to convert the dichotomous responses into a continuous variable. This technique considers between-individual variation as a random factor and has the advantage of higher statistical power compared to traditional analysis using a two-level approach (Moscatelli et al., 2012).

3.3.2 Between groups comparison of contrast

There was a significant effect of contrast ($\beta = 29.141; t(402) = 9.076; p < 0.001$), indicating that the number of correct responses increased with increasing contrast. There was no significant effect of group ($\beta = -0.391; t(402) = -1.274, p = 0.20$), but there was a significant group by contrast interaction ($\beta = 12.313; t(402) = 2.102; p = 0.0036$). Performance increased more quickly with contrast for the migraine-with-aura group, indicating greater contrast sensitivity. This difference was summarised by calculating a 75% threshold from the estimated intercept and slope for each group. Confidence limits were then calculated using the bootstrap method proposed by (Moscatelli et al., 2012). For each of 1000 bootstrap samples, the estimated slope and intercept as well as covariance of these parameters were used to create the intercept and slope estimates for a population of simulated observers of the same size as each group. For each simulated observer, the probit linking function was used to specify the expected proportion of correct responses at each contrast, and a simulated psychometric function was generated by random sampling of a binomial distribution. A group 75% threshold was then calculated using the same generalised linear mixed effects model as was used for the true data. This process was used to create 1000 simulated thresholds, from which the standard deviation of the threshold estimates could be calculated. The thresholds are plotted in Figure 3.3a, and show the increased contrast sensitivity for the migraine-with-aura group in comparison with the control group. From, the bootstrap samples, the proportion of trials on which the threshold was lower for the migraine group (96.9) was significantly higher than chance ($p < 0.001$), assessed using a binomial test.

3.3.3 Between groups comparison of flanker facilitation

Relative 75% thresholds for stimuli with flankers are plotted as a function of the flanker distance in Figure 3.3b. Thresholds were normalised independently for each condition (as established by (Polat & Sagi, 1993)), by dividing each threshold by that estimated in the absence of flankers. These results therefore show the change in thresholds induced by the flankers, with values greater than 1 indicating inhibition and values less than 1 indicating facilitation. For both groups, the presence of flankers tended to reduce thresholds, with the degree of facilitation decreasing as the distance of the flankers from the target increased. This facilitation was analysed using a generalised linear mixed effects model with a probit linking function. Contrast, flanker distance and participant group were used as predictors. Three interaction terms were also included in the model. Contrast-by-flanker distance was included to test whether the slope of the psychometric function was affected by the flanker distance. Group-by-flanker distance was included to determine whether the flanker effect differed between the two groups. Finally, group-by-flanker distance-by-contrast was included to determine whether any effects of the presence of flankers on the shape of the psychometric function differed between the two groups.

Performance increased with increasing contrast (β = 31.46; $t(2428)$ = 9.912; p < 0.001), and decreased with increasing flanker distance (β = −0.05807; $t(2428)$ = −3.680; p < 0.001). There was no main effect of group ($\beta = -0.02525; t(2428) = -0.1716; p = 0.8638$),

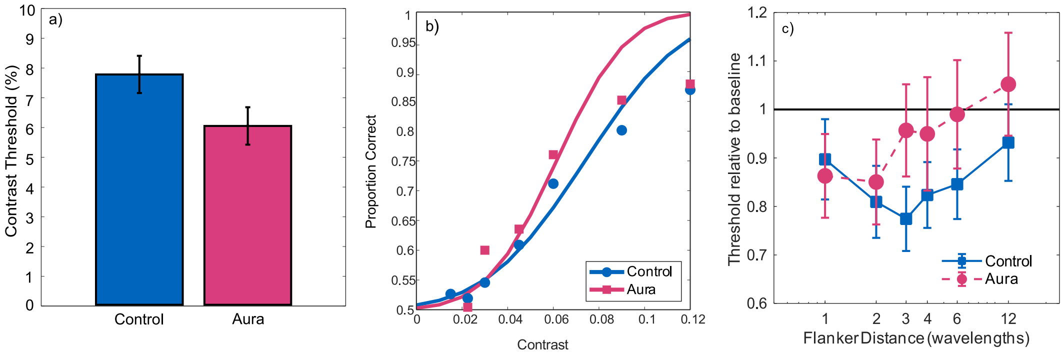


Figure. 3.3 (**a**) Contrast detection thresholds (75% correct) for aura and control groups for Gabor targets in the absence of flankers; (**b**) Interaction in the psychometric functions for the improvement in contrast sensitivity in the baseline (no flankers) condition. Performance increased more quickly for the the migraine-with-aura group (indicated by the steeper slope) which suggests greater contrast sensitivity (authors acknowledge that the psychometric fit is less than ideal) ; (**c**) contrast detection thresholds (75% correct) for stimuli with flankers present. Thresholds are plotted relative to the condition's baseline (marked by the horizontal black line). Normalisation was undertaken independently for the control and migraine groups. Error bars indicate ±1 standard deviation.)

and there were no significant distance-by-contrast ($\beta = 0.2389; t(2428)0.7820; p = 0.434$), group-by-flanker distance (β = 0.02017; t(2428) = 0.8400; p = 0.4010) or group-by-flanker distance-by-contrast (β = $-$0.4107; t(2428) = $-$1.082; p = 0.279) effects. These results show a facilitatory effect of the flankers, which decreased with increasing distance, and that the degree of facilitation did not differ between the two groups.

Performance increased more quickly with contrast for the migraine-with-aura group, indicating greater contrast sensitivity.

3.4 Discussion

The purpose of this study was to directly investigate the lateral interactions in individuals with migraine with aura, to determine whether the deficits experienced were consistent with a reduced inhibitory effect. We found no evidence of differences in inhibitory or excitatory lateral interactions in those with migraine compared to the control group, consistent with previous findings (Shepherd et al., 2011). This does not support the hypothesis that cortical hyperexcitation is due to a lack of inhibition between units. A lack of inhibition would be denoted by a reduced effect of the flankers and a flanker-by-contrast interaction between the migraine-with-aura and control groups. The advantage of the current study is that this has been demonstrated for Gabor stimuli containing a narrow range of spatial frequencies and the same orientation as the target, in comparison with the solid circle and annulus stimuli used by (Shepherd et al., 2011). Thereby controlling for stimuli of the same orientation, and potential confounding visual processes (Baker & Vilidaite, 2014; Polat & Sagi, 1993). Aurora and Wilkinson (2007) suggested that colinear facilitation might be abnormally strong in migraine. While there was no evidence to support this hypothesis in the current study, our results did show that individuals with migraine-with-aura exhibited increased performance on the baseline contrast detection task.

3.4.1 Contrast Sensitivity

One of the most basic ways in which visual performance might differ interictally between migraine and control groups is in contrast sensitivity. While there is some evidence that contrast sensitivity is reduced in individuals with migraine (Braunitzer et al., 2010; Mendes et al., 2011; Shepherd et al., 2012, 2013), the findings are inconsistent (O'Hare & Hibbard, 2016). Some studies have found evidence for poorer contrast sensitivity in migraine when testing with large (Khalil et al., 2011) or peripherally presented stimuli, but not when testing with centrally presented stimuli (McKendrick & Sampson, 2009). There is also evidence for reduced performance in migraine compared to controls for low-spatial-frequency stimuli but not high-spatial-frequency stimuli (Benedek et al., 2002). However, studies using small, briefly presented stimuli, similar to those used here, have found contrast sensitivity in migraine to be as good as controls (Wagner et al., 2010) or even better (Shepherd, 2000). This discrepancy may be related to the variation in sensitivity found across the visual field in people with migraine (Mckendrick & Badcock, 2003) and the variety of psychophysical techniques used across studies (O'Hare & Hibbard, 2016). Increased contrast sensitivity, as found in the current study, is consistent with the idea of hyperexcitation in migraine (Palmer et al., 2000). This is inconsistent with other studies demonstrating reduced-contrast sensitivity in those with migraine, using the Cambridge Low Contrast Gratings (e.g., (Shepherd, 2000; Singh & Shepherd, 2016)). It has been suggested that the reduced contrast sensitivity in migraine shown in some studies can be accounted for by increased internal noise levels in the visual system (Wagner et al., 2010). The Cambridge Low Contrast Gratings are much larger stimuli than the Gabor patches used in the current study, and as there is variation across the visual field in migraine (Mckendrick & Badcock, 2003), this may account for differences between previous research and the current findings.

3.4.2 Inhibition and Lateral Interactions

Many aspects of low-level visual processing rely on excitatory and inhibitory lateral interactions between neurons, for example, visual search, border segmentation, and defining contours (Zhaoping, 2014). Inhibition has been proposed to mediate the strength of sensory activation within the primary visual cortex (V1). Because populations of neurons are sensitive to different ranges of spatial frequencies, channels only respond when the when the frequency of the incoming sensory information matches a channel's frequency tuning. Channels whose frequency tuning does not correspond to the incoming sensory information do not respond. In addition to these direct excitatory responses to external stimulation, the channels also receive signals from neighbouring neural populations that code for similar spatial frequencies or orientations (Polat & Sagi, 1993). These interactions may be different in migraine, potentially accounting for the deficit in performance in some visual detection tasks (O'Hare & Hibbard, 2016). However, the behavioural evidence is inconsistent. For example, no difference in inhibitory interactions in migraine was found for a contrast masking task, in which observers were required to detect a rapidly presented target on a high-contrast mask (McColl & Wilkinson, 2000).

Furthermore, while Battista et al. (2011) found differences in inhibitory interactions between migraine and control groups using the Chubb illusion (Chubb et al., 1989), this was only the case for moving stimuli, and increased, rather than reduced, inhibition was found. Both groups performed at the same level when static stimuli were presented. McKendrick et al. (2017) showed increased surround suppression in migraine using a drifting version of the Chubb illusion, consistent with a reduced level of short-range inhibition measured interictically using visually evoked potentials (Coppola et al., 2013). In addition, there were effects of the migraine cycle—specifically less surround suppression in the migraine group 2 days before the onset of the attack. However, the reduced performance for moving stimuli may not be completely accounted for by reduced lateral inhibition. For instance, evidence suggests that motion after-effects are enhanced in migraine (Shepherd, 2001). The motion after-effect is a visual illusion of motion observed after viewing moving stimuli for an extended period. After adapting to motion in one direction, observers perceive motion in the opposite direction whilst viewing a stationary stimulus (Mather et al., 1998). It has been argued that increased motion after-effects in migraine can best be explained by reduced excitatory intracortical interactions (Shepherd, 2001, 2006; Singh & Shepherd, 2016). If this is the case, we might have predicted a decreased flanker facilitation effect in migraine-with-aura ; however, this was not found in the current study. While no significant difference was found, the curve for the migraine with aura is shifted leftwards relative to the control group Figure 3.3 (b). This shift would not represent a change in the strength of the facilitation effect but in the spatial area over which it operates. In this case, a significant shift would have represented a reduction in the spatial extent of facilitatory connections.

In contrast, the physiological evidence for reduced levels of the inhibitory neurotransmitter in individuals with migraine-with-aura is robust (Bridge et al., 2015). Because reduced GABA is argued to disrupt the excitatory and inhibitory activity in the early visual areas (Bridge et al., 2015), it has been predicted that individuals with migraine should perform worse in tasks such as orientation discrimination, that require strong local inhibitory interactions. Sharpening of the orientation tuning for individual receptive cells depends on lateral inhibition (Seriès et al., 2004), and orientation sensitivity is known to correlate with GABA concentration in the visual cortex (Edden et al., 2009). Tibber et al. (2006) tested this and found as predicted that sensitivity to orientation differences was reduced in migraine. However, the orientation deficit was only found for oblique, not cardinal, orientations; therefore the authors ascribed this finding to difficulties in pooling inhibitory/excitatory interactions, rather than impaired lateral inhibition alone. Additionally, Wilkinson and Crotogino (2000) showed a trend for poorer orientation selectivity in those with migraine; however, as this was not statistically significant, the authors emphasise that this is inconclusive evidence for reduced lateral inhibition in migraine.

In this study, we predicted that should individuals with migraine experience a difference in the balance between inhibitory and excitatory interactions in V1, then the pattern of inhibition and facilitation in this lateral masking task would differ from that of controls. Our results found no difference between migraine and control groups. Our study adds to the literature that suggests lateral inhibitory processes may not be exclusively responsible for the perceptual differences experienced in those with migraine. While reduced GABA levels are argued to cause disruption in excitatory and inhibitory activity within visual areas and are correlated with cortical excitability (Bridge et al., 2015), this does not appear to impact lateral interactions in behavioural tasks in a consistent and reliable manner. Wagner et al. (2010) proposed that reduced cortical suppression may be reduced through GABAergic inputs, and this may also impact the efficiency with which individuals with migraine can exclude noise.

3.4.3 Global Processes and External Noise

Given the lack of any differences in lateral interactions from our study, the alternative explanation for impaired performance in migraine is the failure to ignore external noise when performing global integration tasks (Tibber et al., 2014; Wagner et al., 2013, 2010). Global processes have been studied in migraine, using both global form (Ditchfield et al., 2006; McKendrick et al., 2006) and global motion tasks (Shepherd et al., 2012; Tibber et al., 2014). Global form perception in migraine has been assessed with Glass patterns (Ditchfield et al., 2006; McKendrick et al., 2006). Glass patterns typically consist of pairs of individual elements aligned to show a global form structure. Noise was introduced to these stimuli by varying the orientation of the elements relative to each other. These studies demonstrated higher thresholds (poorer performance) in migraine groups for Glass pattern detection. This result might be taken as evidence that those with migraine are poorer at contour integration, a process that depends on the lateral interactions in the association field (e.g., (Field et al., 1993; Hess et al., 2003)). However, no evidence for differences in lateral interactions was found in the current study. The difference between the Glass patterns and the current study stimulus is the inclusion of additional noise elements in the former. In the Glass pattern task, the aim is to detect the coherent orientation of the pattern, with noise typically added by increasing the proportion of randomly oriented “noise” elements. This involves at least two processes, the contour integration, and also the discrimination of the elements of coherent orientation against the noise elements.

Analogously, motion coherence is typically studied using random-dot kinematograms. In a typical stimulus, there are a proportion of signal dots that move in one direction while the remainder move randomly. Difficulty is manipulated by varying the number of signal dots; as it is not possible to identify the motion from a single dot, the mean direction is inferred by segregating out the noise dots and integrating the coherently moving signal dots (Tibber et al., 2014). People with migraine have shown a deficit in judging the direction of motion in these kinds of tasks (Antal et al., 2005; Ditchfield et al., 2006; Shepherd et al., 2012; Tibber et al., 2014). In these tasks, local motion needs to be detected, the coherent direction integrated, and the randomly moving “noise" elements segregated from the coherently moving stimuli (Dakin et al., 2005a). Motion coherence can also be studied using a paradigm of equivalent noise. In this method, rather than assigning signal and noise dots, all dots contribute equally to the signal, and the difficulty of the task is increased by varying the distribution of the direction of motion of the dots (see (Tibber et al., 2014)). Observers with migraine were found to perform normally on equivalent noise motion coherence tasks. This suggests that rather than a deficit in pooling signals, people with migraine may have difficulty segmenting signals from randomly moving “noise” elements in the more typical global motion task. This suggests that deficits in migraine may not be due to lateral interactions needed to integrate the coherent motion but are from the exclusion of noise from the signal. The inability to exclude noise from signals has been suggested as the reason for the differences seen in migraine in masking studies (Wagner et al., 2010).

There are also temporal factors that affect differences in lateral interactions in migraine studies. Typically, those studies that have identified robust differences between migraine and control groups are those using moving stimuli. For example robust differences have been found in the perception of drifting gratings (such as the Chubb illusion (Battista et al., 2011; McKendrick et al., 2017)), prolonged motion after-effects (Shepherd, 2001, 2006; Singh & Shepherd, 2016), and poorer performance in tasks involving the discrimination of motion against noise backgrounds (Shepherd et al., 2012; Tibber et al., 2014). Increased masking in migraine is found at all stimulus onset asynchronies, but the exact timing of the target relative to the mask makes a difference—the masking effect is greatest for 60–100 ms intervals (Wagner et al., 2013). The authors modelled this effectively by increasing both the variable representing gain control instead of the post-gain-control noise variable in their perceptual template model. The post-gain-control noise variable relates to the shape-encoding mechanism in the model. GABA is related to the gain control of the visual system, and although GABA is associated with orientation specificity, it might not be directly relevant to this higher-level shape mechanism (Katzner et al., 2011).

3.5 Conclusions

Contrast sensitivity was increased in those with migraine with aura, which was consistent with cortical hyperresponsivity. This is contrary to other measures of contrast sensitivity in migraine. The difference between the current study and previous findings could be related to the stimulus size, as performance varied across the visual field in migraine, and orientation-specific deficits tend to be for lower-spatial-frequency stimuli. This study does not support the hypothesis of altered lateral interactions in those with migraine-with-aura. It is suggested that the overall patterns of poorer performance in those with migraine in global tasks could be due to a reduced ability to exclude noise, rather than differences in lateral interactions per se.

Acknowledgements

Partial data collection was undertaken as part of a UROP internship, however all writing, analysis and remaining data were collected during the PhD

Chapter 4

Spatial frequency tuning of perceptual learning and transfer in global motion

4.1 Introduction

Perceptual learning has attracted much attention as a potential tool to aid recovery of lost visual function for clinical populations (Campana & Maniglia, 2015). The success of perceptual training in amblyopia (Hess et al., 2003; Levi & Li, 2009; Polat, 2009), presbyopia (Polat, 2009) and cortical damage (Huxlin et al., 2009; Sahraie et al., 2006; Trevethan et al., 2012)may contradict the position that sensory development is restricted to a critical period early in life (Sagi, 2011; Seitz & Watanabe, 2005) and that the visual system is hard-wired in mature systems (Hubel & Wiesel, 1970). While it has repeatedly been established that training can improve perceptual abilities (Fahle, 2005), these benefits tend to be highly specific for both the perceptual features of the stimuli (Ahissar & Hochstein, 1997; Fiorentini

& Berardi, 1980; Karni & Sagi, 1991) and the behavioural task used in training (Ball & Sekuler, 1982). This specificity severely limits the effectiveness of perceptual learning as a general therapeutic tool. Resolving the conditions under which learning is tied to the features and tasks used in training, and how much it can generalise to new tasks and stimuli, is imperative for understanding the mechanisms of perceptual learning (Fahle, 2005; Lu et al., 2011).

4.1.1 The Reverse Hierarchy Theory

The Reverse Hierarchy Theory an early theoretical model of perceptual learning (as introduced in Chapter 1), suggests that the key to understanding specificity and transfer lies in the hierarchy of processing and the feedforward and feedback connections between them (Hochstein & Ahissar, 2002). This theory predicts that an increase in sensitivity at lower cortical areas would be as a result of the feedback connections from higher cortical areas. Thus, paradigms that incorporate the higher-stage (global) aspects of perception to evaluate the transfer of learning may shed light on the nature of these learning mechanisms.

The visual system is known to operate in a hierarchical manner and motion signals are initially processed in V1 (Livingstone & Hubel, 1988; Movshon & Newsome, 1996). where the receptive fields are small and only respond to a restricted area of the visual field (Hubel & Wiesel, 1962). Since each receptive field is highly tuned, the response obtained from each cell only contains one-dimension of motion for the stimulus (Adelson & Movshon, 1982; Amano et al., 2009) (such as direction, location or orientation). However, in order to process complex motion where signals are spatially, temporally and directionally incoherent, the restricted local signals require pooling and integrating in order to produce a multidimensional or ‘global’ percept. However, as perceptual beings, we are unaware of the highly specific individual frequency components of complicated stimuli (Ahissar et al., 2009). The gap between the highly tuned frequencies and our holistic perception is facilitated via local and global processing (Ahissar et al., 2009). A number of areas are known to play a role in processing global motion, including V3A, V3 and V5. The receptive fields in these visual areas are larger and more broadly tuned than those in V1 (Felleman & Van Essen, 1987), and sum the responses of a set of V1 neurons across space, time, orientation and spatial frequency (Simoncelli & Heeger, 1998).

In the case of motion coherence, improvements would for example be predicted to be dependent on the spatial frequency tuning of the higher-stage motion detectors (Ahissar & Hochstein, 2004). The broader frequency tuning found at this level, in comparison with the local stage of motion processing (Bex & Dakin, 2002), is characteristic of the generalising properties of higher-level processing (Hochstein & Ahissar, 2002). Consistent with the predictions of the Reverse Hierarchy Theory, observers trained on a global motion coherence task, and reported the expected improvement in performance on this task (Levi et al., 2015). However, observers also improved in their contrast sensitivity measured using sinusoidal gratings. No training had been provided on the task or the stimuli. This result demonstrates the potential for generalisation of learning achieved through training on global motion tasks. While, improvement was restricted to low spatial frequency drifting sinusoidal gratings, improvement was found in untrained retinal locations and the untrained eye (Levi et al., 2015). Backward projections play an important role in the perception of global form and motion, and their modification represents a theoretically plausible important component of perceptual learning. Transcranial Magnetic Stimulation (TMS) has been used to stimulate the re-entrant connections from V5 to V1, enhancing the perception of coherent motion (Romei et al., 2016a). Using a novel paired cortico-cortical TMS protocol (ccPAS) to induce Hebbian plasticity, observers' thresholds for motion detection were reduced when the feedback connections from V5 to V1 were stimulated. However, the improvement in perception was critically dependent on the timing and direction of stimulation. There was no change when the feedforward connections from V1 to V5 were stimulated (Romei et al., 2016b). This also suggests that these re-entrant connections are malleable (Romei et al., 2016b). Using a similar method, a direction-selective improvement was induced by pairing subthreshold stimulation with the simultaneous presentation of direction-specific moving stimuli (Chiappini et al., 2018). This provides additional support for the accumulating evidence that the re-entrant connections from direction-tuned neurons in areas such as V5 play a role in perceptual learning in global motion coherence tasks. Sillito et al. (2006) suggest that the “iterative interaction" between the two stages could account for the selectivity at both levels. Backward projections are a theoretically plausible route for perceptual learning, and are consistent with the predictions from the Reverse Hierarchy Theory. Overall this suggests a theoretically important role for top-down information in obtaining transfer from global motion (Rockland & Knutson, 2000).

This study aims to identify the relative location of one of the mechanisms involved when learning direction discrimination for a global motion coherence task. We do this by evaluating the spatial frequency tuning of improvements in performance for trained and untrained tasks. Given the differences in spatial and temporal frequency tuning of processing between the local and global levels (DeValois et al., 1982; Lui et al., 2007; Newsome et al., 1988), measuring the tuning of training for these dimensions allows us to understand the role of each level in global motion learning. Specifically, if the changes in neural weightings that allow learning involve the global level, we would expect it to show the broad, relatively low-pass frequency tuning that characterise this processing stage.

4.1.2 Experiment 1

Prior to collecting the data for our main study, we also questioned whether trial-by-trial feedback was a requirement for learning. Therefore, we first investigated the necessity of feedback for perceptual learning to occur for our specific stimuli. The specific nature of feedback and its role in perceptual learning is unclear, and external feedback has been shown to improve learning and increase efficiency (Dobres & Watanabe, 2012; Herzog & Fahle, 1997). However, some studies have found that learning occurs without external feedback (Ball & Sekuler, 1982; Fahle, 2005; Liu et al., 2012a; Petrov et al., 2006a; Shiu & Pashler, 1992; Vaina et al., 1998). Recently, it has been found that interleaving high accuracy (easy) trials and low accuracy (difficult) trials resulted in perceptual learning without the need for feedback, even on difficult trials (Liu et al., 2012a). Based on these results we predicted that we should find learning in both conditions, as long as easy and difficult trials were interleaved. As detailed in the following sections, our study found that learning did not occur in the condition where no feedback was provided, even with easy trials. Learning only occurred for the with-feedback condition. With this in mind our design for experiment 2 included feedback during training, but no feedback when testing.

4.1.3 Experiment 2

The purpose of this study was to extend the pre- and post-training measures captured by Levi et al. (2015). Participants were trained on one of three motion coherence stimuli where individual elements were comprised of either broadband (referred to as broad) Gaussian blobs, or low- or high-frequency random-dot Gabors. Following training, they were assessed for improvements in global motion and contrast sensitivity for static patterns. This assessment was done for a range of spatial frequency profiles. Thus, as well as testing for transfer to contrast sensitivity, we included measures of transfer to the trained and untrained global motion frequencies.

Transfer from global motion could involve a combination of reweighting that includes the global motion mechanisms in higher areas such as V5, and the re-entrant connections from global mechanisms to local mechanisms. Differences in the spatial frequency tuning at the local and global stage mean that these models make different predictions for the transfer of learning across spatial frequency, and different psychophysical tasks (Figure 4.1). These will be addressed in detail in the following section.

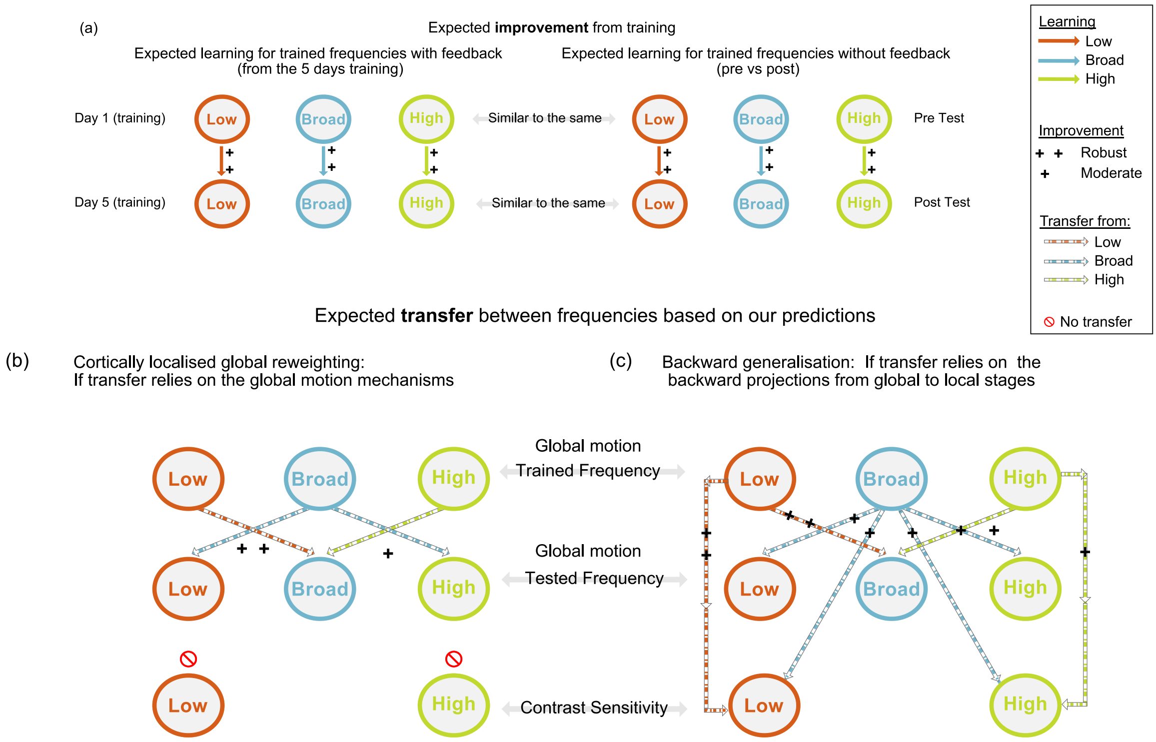


Figure. 4.1 (a) **Expected improvement from training:** We would expect robust improvement on frequency specific motion stimuli over the 5 days training, and for this to be reflected at the post-assessment as an improvement compared to the pre-assessment results. Predictions of transfer (b) **Cortically localised global reweighting:** In this case we may predict that transfer from frequency specific global motion would reflect the broad, approximately low-pass spatial frequency tuning of global motion detectors in areas such as V5 and V3/V3A. There would thus be most transfer when both training and test stimuli contained low spatial frequency components(c) **Backward generalisation:** Should transfer occur as a result of the backward projections from V5 to V1 (as predicted by the Reverse Hierarchy Theory (Ahissar & Hochstein, 2004)) we would therefore also expect robust transfer from low to low frequency stimuli; modest transfer from low to broad, and broad to low and broad to broad, with an unlikely but possible transfer from high to broad, and broad to high. Importantly, should there be transfer to contrast sensitivity we would expect this to reflect the low-pass spatial frequency tuning of global areas.

4.2 Predictions

4.2.1 Cortically localised global reweighting: If transfer relies on global motion mechanisms

If transfer is restricted to the cortical location of processing, we would predict that any improvements after training using global motion coherence tasks would reflect the more broadly tuned frequency tuning of global motion detectors in areas such as V5 and V3/V3A (Amano et al., 2009; Bex & Dakin, 2002; Felleman & Van Essen, 1987; Henriksson et al., 2008). Thus, transfer would be most likely when the training and test stimuli contain low spatial frequency components that mirror the preference of global motion detectors (robust transfer is predicted from low frequency to low frequency conditions; modest transfer from low to broad frequency stimuli, and from broad to low and broad stimuli, and an unlikely but possible transfer from broad to high and high to broad frequency conditions.). No transfer to contrast sensitivity would be predicted, since learning would result in changes in weightings between direction-tuned global motion mechanisms, at the higher-level decision stages only (See Figure 4.1(b))..

4.2.2 Backward generalisation: If transfer relies on feedback projections from global to local stages of processing

The distinctive difference between the two predictions is the transfer of learning to the untrained, locally processed contrast sensitivity task. Improvement in contrast sensitivity performance requires extensive training (Sowden et al., 2002; Yu et al., 2004), and is highly tuned to spatial frequency, specific for the trained eye and retinal location of training (Adini et al., 2004; Kuai et al., 2005; Sowden et al., 2002; Xiao et al., 2008; Yu et al., 2004; Zhang et al., 2010, 2008). Should improvement occur for contrast sensitivity this would suggest a role for the reentrant feedback connections, and support a backwards generalisation of transfer as predicted by the Reverse Hierarchy Theory (Ahissar & Hochstein, 2004). In this case improvement would be expected to reflect the tuning properties of the global motion processing areas. Based on the attenuation of high spatial frequencies from global motion areas (Amano et al., 2009), we would expect most transfer when both training and test stimuli contained low spatial frequency components (robust transfer from low to low frequency stimuli; modest transfer from low to broad, and broad to low and broad to broad, with unlikely transfer from high to broad, or broad to high frequency). (Figure 4.1 (c)).

We would also predict transfer to contrast sensitivity, and for this to show the same low-pass spatial frequency tuning, consistent with the transfer found by Levi et al. (2015). This would be evidenced by improved performance for low spatial frequency contrast, but no change for high spatial frequency contrast.

The distinctive difference between the predictions is the transfer to contrast sensitivity. Should transfer occur to contrast sensitivity, that is tuned to the frequency of training, this would suggest a role for the reentrant feedback connections, an support a backwards generalisation of transfer.

4.3 Methods and Materials:

4.3.1 Participants

24 observers for Experiment 1 and 30 (new) observers for Experiment 2 were randomly and evenly assigned into groups. For experiment 1 there was a feedback and a No-Feedback group. For experiment 2, groups were categorised by the spatial frequency of training (broad, low, high). All observers were employees or students from the University of Essex and self declared as having normal or corrected-to-normal vision. All work was carried out in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki). The study procedures were approved by the University of Essex Ethics Committee (JA1601). All observers gave informed written consent and were either paid or received course credit for their participation.

4.3.2 Experiment 1:

Stimuli were generated and presented with Matlab 2015a using the Psychophysics Toolbox extensions (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997). The broadband motion coherence stimuli were presented using a 2.7 Ghz iMac running OSX 10.9.5. The 27" monitor had a display resolution of 2560 x 1440 pixels and a 60 Hz refresh rate. Viewing distance was 450mm, the stimuli subtended a visual angle of 66.8°, and one pixel subtended 1.77 arc minutes. Motion stimuli contained 100 Gaussian elements, each with a standard deviation of 6.8 arc minutes (see Figure 4.2a). Motion stimuli were based on the task designed by

Williams and Sekuler (1984), where the direction of motion of each dot is drawn from a uniform probability distribution (levels of coherence). Elements were presented within a mid-grey rectangular region measuring 17.6° x 17.6° on a mid-grey background. Elements were presented for 1 second, moving 5 pixels per frame and a total distance of 8.8°. Dots wrapped around the edges of the rectangle. The seven levels of coherence (5°; 10°; 15°; 20°; 25°; 60°; 180°) were determined using a pilot study where accuracy across levels confirmed that there was an even distribution of easy trials (85% and above) and difficult trials (65% and below), with the balance around 75%. Levels were presented as two-alternative-forced-choice (2AFC) trials, interleaved in a random order within each session and used the method of constant stimuli (MOCS). There were 40 repetitions of each level and responses were obtained via the left or right arrow key associated with the perceived direction of coherent motion. Feedback (if present) consisted of an auditory beep immediately following each response, a high pitched tone for a correct response (2000Hz for 10ms), and a low pitched tone for an incorrect response (200Hz for 40ms).

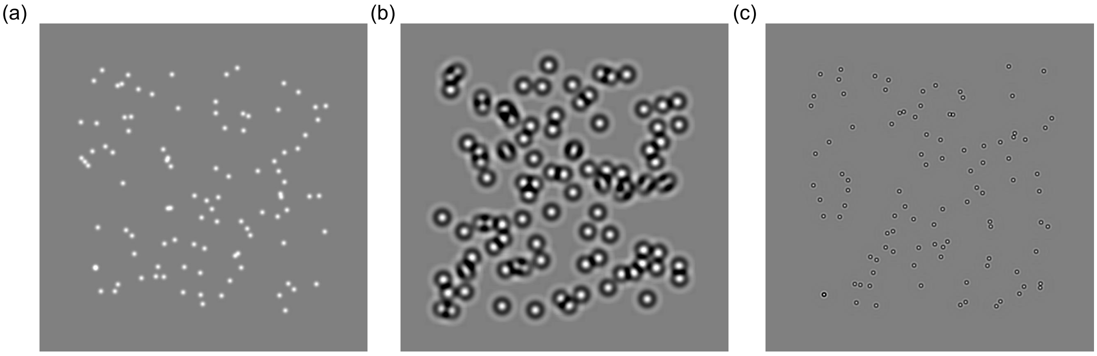


Figure. 4.2 Stimuli consisted of three global motion conditions where individual elements were comprised of either broadband Gaussian blobs, or low- or high-frequency random-dot Gabors). For all stimuli the speed of motion was held constant.

4.3.3 Experiment 2:

A series of baseline measures were completed for motion coherence direction discrimination and contrast sensitivity. Test (Pre and Post) stimuli were presented on a VIEWPixx/3D 23.6 inch monitor with a display resolution of 1920 x 1080 pixels, and a 120 Hz refresh rate, using a Dell Precision T3610 PC running Windows 7. One pixel subtended 1.6 arc minutes and stimuli were viewed for 1000 milliseconds (60 frames per second) from a distance of 570mm. Head position for testing was stabilised using a chin rest. Between these testing sessions, observers undertook five consecutive days of global motion training in one of three spatial frequency groups (broad, low or high). Training stimuli were presented on a 19" monitor with a display resolution of 1980 x 1080 pixels and 60 Hz refresh rate, using a PC running Windows 7. Stimuli sizes and presentation times were adjusted to accommodate for the different specifications. One pixel subtended 1.7 arc minutes. Observers were positioned with a viewing distance of 500mm, which was checked routinely with a measured piece of string. The stimuli were presented for 1000 milliseconds, moving 10 pixels per frame (30 frames per second).

Training Stimuli

Global Motion: Broadband stimuli were the same as previously described. For low-frequency stimuli, the elements were circularly symmetric Gabor patches. The standard deviation of the Gaussian window, σ, was 30.1 arc minutes, and the spatial frequency of luminance modulation, f, was 1 cycle/degree. For each element, the luminance profile was defined as a function of horizontal and vertical position (x,y) as:

$$d(x,y) = \sqrt{(x-x_0)^2 + (y-y_0)^2} \tag{4.1}$$

$$L(d) = \frac{A}{\sigma\sqrt{2\pi}} \exp\frac{-d^2}{2\sigma^2} \cos 2\pi f d \tag{4.2}$$

where x_0 and y_0 is the central position of the element, and A determines its contrast. Elements for the high-frequency stimuli were defined in the same way, but had a standard deviation of 7.48 arc minutes and a spatial frequency, f, of 4 cycles/degree. For all stimuli, the spatial frequency of the elements and the speed of motion were held constant. Initially all elements were uniformly and randomly distributed within a region of 16.6°x 16.6° in the centre of the screen. A central black fixation dot was presented at all times when stimuli were not being displayed. Examples of the stimuli are shown in Figure 4.2 a-c. Motion was created using the method previously detailed, and dots moved a fixed distance of 8.5 arc minutes per frame.

Testing Stimuli (Pre and Post)

1. *Global motion:* Stimuli were identical to those described in the training session, with the following exceptions. The standard deviations of the elements were 6.4 arc minutes (broadband), 28.4 arc minutes (low-frequency) and 7.0 arc minutes (high-frequency).

Stimuli were presented within a mid-grey rectangle measuring 15.9°x 15.9°, each element moved a fixed distance of 8 arc minutes.

2. *Contrast Sensitivity:* Stimuli were Gabor patches, with a spatial frequency of 1 cycle per degree (/°) or 4 cycles/°, presented in the centre of the screen on a mid-grey background, tilted either ± 20° away from vertical. The Gaussian envelope of the Gabor stimulus had a standard deviation of 1.1°. 7 levels of contrast (0.05, 0.1, 0.15, 0.175, 0.2, 0.3, 0.4 % Michelson Contrast) were presented.

4.3.4 Procedure

For each observer, training was undertaken at one spatial frequency only, totalling 420 trials daily for 5 consecutive days. Based on the findings from experiment 1, feedback was provided after each trial. Testing (pre and post) measures were taken for motion coherence (for all frequencies) and contrast sensitivity (high and low spatial frequency). Responses were captured on a DataPixx response box for contrast sensitivity, and left and right arrows on the keyboard for the motion task. The presentation order of trials was randomised for direction and coherence (for global motion), spatial frequency and orientation (for contrast sensitivity). There were 20 repetitions for each of the seven levels, for each condition. Testing was performed in a darkened room, before and after training.

4.4 Results

4.4.1 Statistical Methods

Learning is often measured through monitoring performance at a particular threshold, which is expected to shift the psychometric function leftwards if performance is improved, (see Figure 4.3 (a)). The psychometric function describes performance in terms of accuracy as a function of the strength of the stimulus (Gold et al., 2010), which are usually positively correlated, and performance is expected to reach an asymptotic level at the highest stimulus intensity (Swanson & Birch, 1992). Inspection of the pre- and post-assessment data revealed that in some conditions performance did not reach perfect accuracy, asymptoting at a proportion of correct responses that was less than 1; this resulted in a poor psychometric fit of the observer response data using a GLMM. To accommodate the variability in asymptotic performance, a nonlinear generalised mixed effects model was used (Donovan & Carrasco, 2015; Herrmann et al., 2010; Ling & Carrasco, 2006). The nonlinear regression provides three measures to assess a change in performance over time. Firstly, like the GLMM a leftward shift in the curve indicates an improvement in threshold (Figure 4.3 (a)). An increase in slope indicates an increase in the rate at which performance increases with increase signal level (Figure 4.3 (b)). Finally a change in the asymptote indicates a significant change to the performance at the highest level of stimulus intensity (Figure 4.3 (c)). These changes are independent aspects of the psychometric function fit, and may not necessarily be congruent.

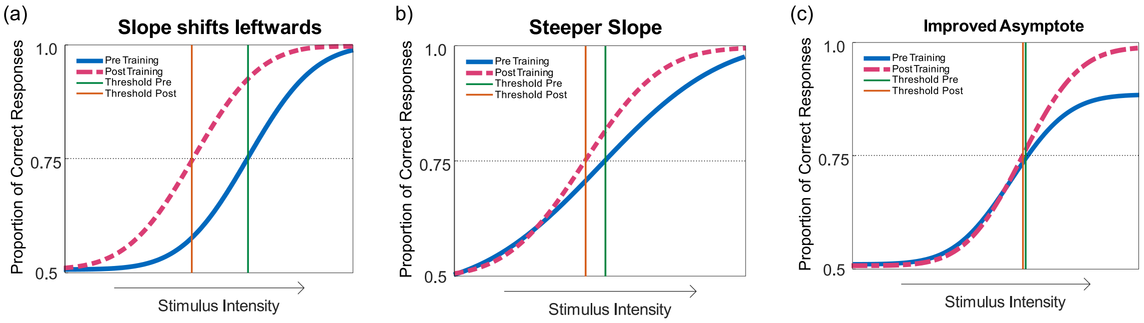


Figure. 4.3 Psychometric functions illustrating the three measures by which the non-linear regression provides evidence of a change. (a) A leftward shift of the function indicates a general increase across stimulus levels, that did not vary as stimulus intensity increased. (b) A steeper slope indicates an increase in the number of correct responses as stimulus intensity increases. (c) An upward shift of the asymptote indicates an increase in performance where stimulus intensity is at its highest (lapse rate). However, in this example there has been no shift of the midpoint.

For example, it is possible to obtain an increase in one measure and a decrease (or no change) in another.

Thus, all analyses were thus conducted using a Nonlinear Mixed Effects Model (NLME) with Matlab (using the *nlme* function) and a logistic link function defined by the following model:

$$p = 0.5 + \frac{(A + A_d(S-1))}{1 + e^{-(K+K_d(S-1))(C-(C_0+C_{0d}(S-1)))}} \tag{4.3}$$

where p is the proportion of correct responses, A determines the asymptotic level of performance, K defines the slope and C_0 defines the midpoint of the function. d is the post-training change in performance, and A_d, K_d and C_{0d} determine the change in asymptote, slope and midpoint, respectively. C is the coherence level, and S is day. Random effects were included for all parameters, and the model was compared to a null model in which only the intercept was free to vary, using the chi-squared statistic to test for a significant improvement in the log likelihood. 95% confidence limits were calculated for the estimated parameter values. These are plotted for each comparison as a nested bar plot within each model fit (reported below). In the case of the contrast sensitivity task, the model with the full random-effects structure did not converge, and a reduced model with random effects for asymptote and slope, and the change in these two parameters between sessions, was used instead.

4.4.2 Feedback and Perceptual Learning: Experiment 1

Results for the feedback and no-feedback conditions are plotted in Figure 4.4. To identify if either trial-by-trial feedback, or a combination of easy and difficult trials in the absence of feedback produces learning, the response data for each condition (Feedback and No Feedback) were analysed independently.

For the No-Feedback Group, there was a reduction in asymptote, a reduction in midpoint and an increase in slope as a result of training. However, from Figure 4.4 it is clear that when feedback was not provided overall improvement in performance is questionable. In contrast, for the Feedback Group, there was an increase in asymptote, a decrease in midpoint, and an increase in slope across the 10 days. All three of these effects indicate an improvement in performance through training. Overall, there was consistent improvement in performance with feedback, but little change when it was not present. These results demonstrate that, for our training procedures, feedback is necessary for robust learning. The role of feedback in perceptual learning remains a complex, unresolved issue (Liu et al., 2012a; Seitz et al., 2006). For the purposes of our study, trial-by-trial feedback was thus provided in all training conditions in the second experiment. The model fits are summarised in Table 4.1.

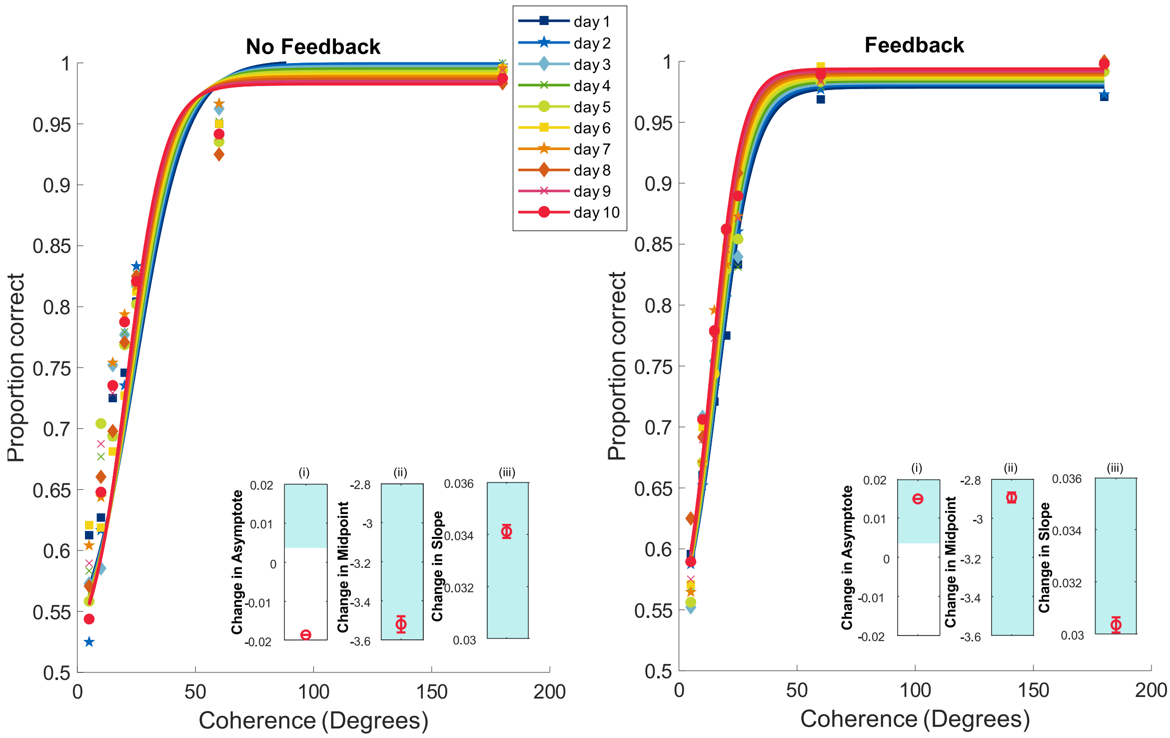


Figure. 4.4 Progress during the ten days training for each feedback condition; (a) No Feedback and (b) with Feedback. Nested plots show the 95% confidence intervals for the change for each measures of interest, namely (i) asymptote, (ii) midpoint and (iii) slope. Results that are exclusively in the shaded region indicate a significant improvement in performance. Results wholly in the unshaded areas show a significant worsening in performance. Results that cross both zones (i.e include zero) show no significant effect of training on that parameter.

Table 4.1 Goodness of fit tests for each model. A mixed-effects model with the intercept as the only free parameter was compared with the model including the full psychometric function. -2 x the difference in log likelihood between the full and null model was calculated, and tested for significance against a chi-squared distribution, with degrees of freedom equal to the difference in the number of parameters for the two models.

Feedback condition	**-2LL**	**ΔDOF**	***p***
Feedback	1528.96	10	<0.0001
No Feedback	1428.24	10	<0.0001

4.4.3 Feedback and perceptual learning: Experiment 2

Predictions

We predicted that should the mechanism of transfer be located at the global motion processing level, and transfer across frequencies would most likely occur for stimuli with frequency properties that reflect the broadband low-pass frequency tuning of global motion detectors. No condition would show transfer to contrast sensitivity, as these are processed in different cortical locations. For global motion, we predicted;

- For the broad frequency trained condition; (i) robust learning on their own trained frequency (ii) moderate transfer to low spatial frequency global motion, and (iii) an unlikely but possibly small improvement to the high frequency condition.
- For the low frequency trained condition; (i) robust learning on their own trained frequency and (ii) robust transfer to broad spatial frequency global motion.
- For the high frequency trained condition we predicted; (i) moderate learning on their own trained frequency and (ii) an unlikely but possibly small improvement to the high frequency condition.

Conversely, should transfer occur as a result of backwards generalisation using the re-entrant connections from global motion processing areas to V1 we would predict that transfer would reflect the pooling of spatial frequencies and attenuation of high spatial frequencies. Specifically,:

- For the broad frequency trained condition; (i) robust learning on their own trained frequency (ii) moderate transfer to low spatial frequency global motion, (iii) an unlikely but possible small improvement in high spatial frequency global motion and (iv) a moderate transfer to low spatial frequency contrast sensitivity.
- For the low frequency trained condition; (i) robust learning on their own trained frequency and (ii) robust transfer to broad spatial frequency global motion, and (iii) robust transfer to low spatial frequency contrast sensitivity.
- For the high frequency trained condition we predicted; (i) moderate learning on their own trained frequency and (ii) an unlikely but possible small improvement for the broad frequency condition, and (iii) no transfer to any frequency of contrast sensitivity.

Training Results

Results for the five days of training, with feedback, are shown Figure 4.5. For the group trained with low-frequency stimuli, there was an increase in asymptote, an increase in midpoint, and an increase in slope. For the group trained with broad-frequency stimuli, there was an increase in asymptote, a reduction in midpoint, and a change in slope. For the group trained with high-frequency stimuli, there was a reduction in asymptote, a reduction in midpoint, and an increase in slope. As can be seen in Figure 4.5, performance on the motion coherence task improved across the training session for all three groups, showing perceptual learning for all spatial frequencies. The model fits are summarised in Table 5.3.

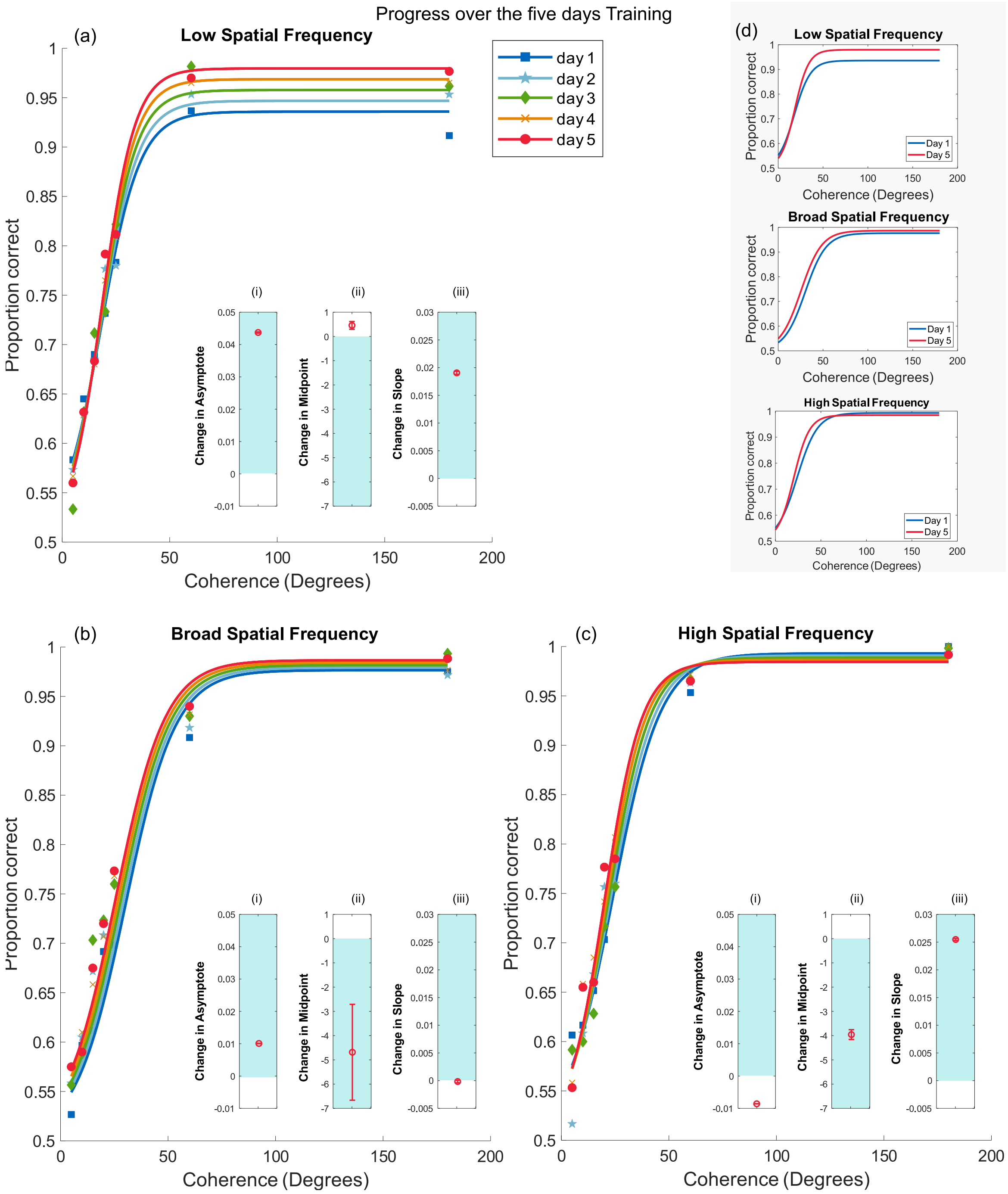


Figure. 4.5 Progress during the five days training for each frequency trained group; (a) Low, (b) Broad and (c) High. Nested plots show the 95% confidence intervals for the change for each measure of interest. Namely (i) asymptote, (ii) midpoint and (iii) slope. Any results that are exclusively in the shaded are indicate a significant positive change (improvement). Results in the unshaded areas show a a significant negative change (worsening performance). Finally results that cross both zones (i.e include Zero) show no significant difference. (d) Fitted curves for the first and last training session are plotted to clarify the effects on training for each frequency.

Table 4.2 Goodness of fit tests for each model. A mixed-effects model with the intercept as the only free parameter was compared with the model including the full psychometric function. -2 x the difference in log likelihood between the full and null model was calculated, and tested for significance against a chi-squared distribution, with degrees of freedom equal to the difference in the number of parameters for the two models.

Training Frequency	**-2LL**	**ΔDOF**	***p***
Low	641.53	10	<0.0001
Broad	618.32	10	<0.0001
High	765.47	10	<0.0001

Pre- and Post-Test Results for Motion Coherence

a) Low-Frequency Training:

Results for the group trained with low-frequency stimuli are shown in Figure 4.6(a). The model fits are summarised in Table 4.3. When tested with low-frequency stimuli, there was an increase in asymptote and a reduction in midpoint, but a reduction in slope. Overall, the results show an improvement in performance, particularly at high coherence levels. When tested with broad-frequency stimuli, there was an increase in asymptote, and an increase in slope, but no change in midpoint. Again, these results show an improvement in performance, most notably at high coherence levels. Finally, when tested with high-frequency stimuli, there was a reduction in asymptote, a reduction and slope and no change in midpoint. There was no improvement in the perception of motion coherence in this case, rather a slight reduction in performance.

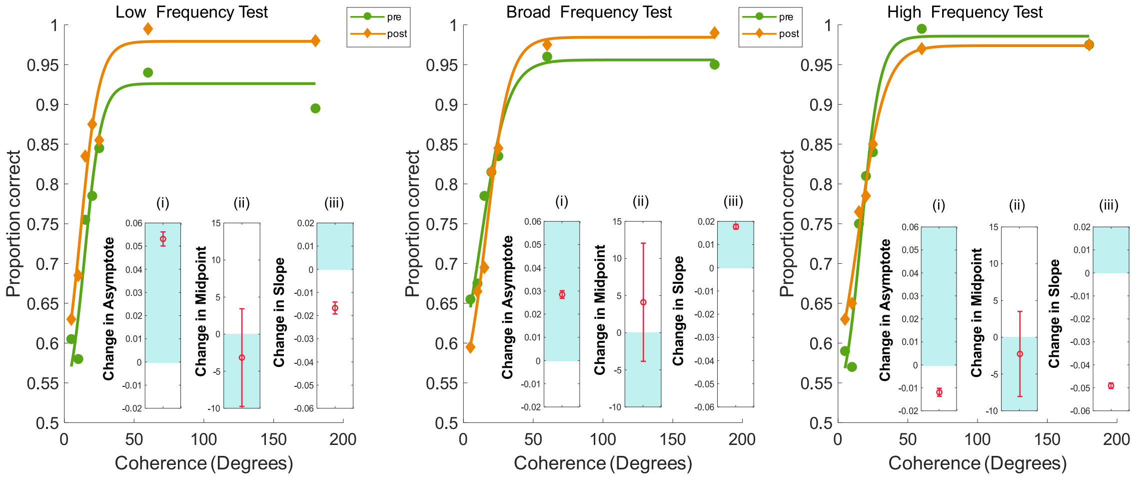
b) Broad-Frequency Training:

Results for the group trained with broad-frequency stimuli are shown in Figure 4.6(b). The model fits are summarised in Table 4.3. When tested with low-frequency stimuli, there was an increase in slope and asymptote but no change in midpoint. As can be seen in Figure 4.6(b), performance improved following training for low contrast test stimuli. in the way of training. When tests with broad-frequency stimuli, there was a reduction in asymptote, no change in midpoint, and an increase in slope. Unexpectedly, these results show a reduction in performance at high coherence levels. Finally, when tested with high-frequency stimuli, there was a reduction in asymptote, an increase in slope and no change in midpoint. Again, these results show an unexpected reduction in performance at high coherence levels after training.

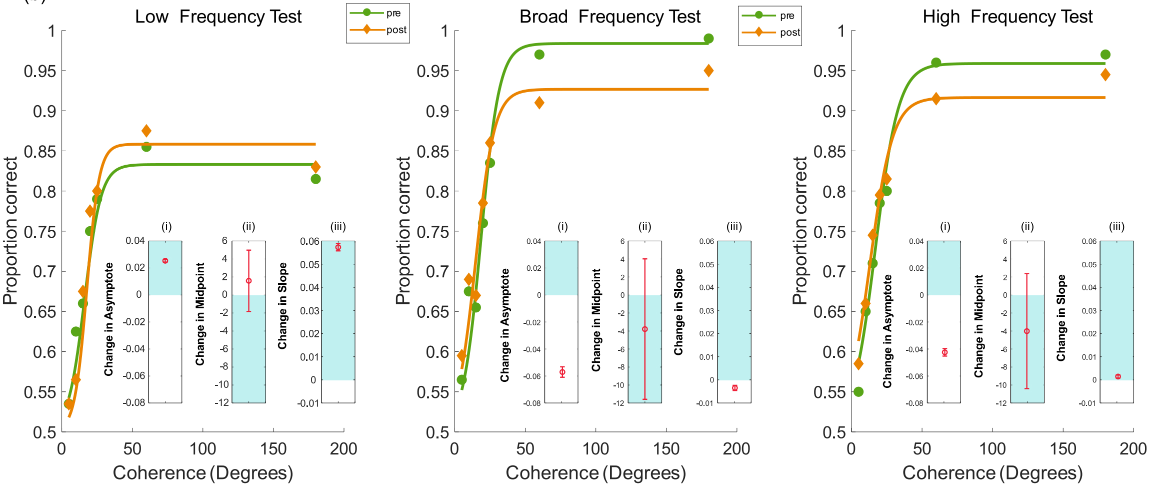
c) High-Frequency Training

Results for the group trained with high-frequency stimuli are shown in Figure 4.6(c). The model fits are summarised in Table 4.3. When tested with low-frequency stimuli, there was a reduction in slope but no change in asymptote or midpoint. There is little evidence for any change in performance following training. When tested with broad-frequency stimuli, there was a reduction in asymptote, no change in midpoint, and an increase in slope. Again, there is little evidence for any change in performance following training. Finally, when tested with high-frequency stimuli, there was a reduction in asymptote and slope, but no change in midpoint. As with the other two test frequencies, there is little evidence for any change in performance.

(a) **Low Spatial Frequency Trained - Global Motion Coherence**



(b) **Broad Frequency Trained Global Motion Coherence**



(c) **High Spatial Frequency Trained Global Motion Coherence**

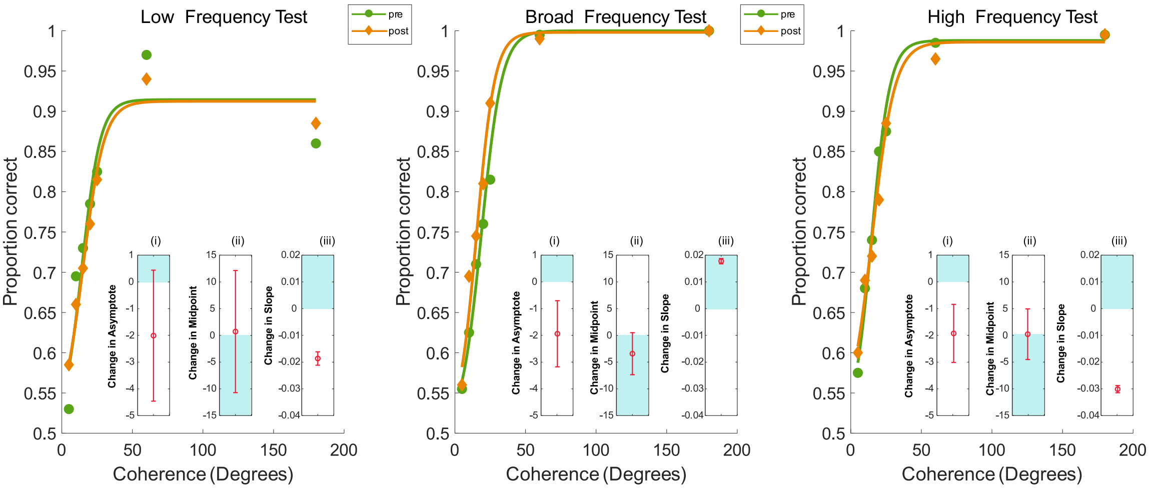


Figure. 4.6 Pre and Post Test Motion coherence results following after the five days training for each frequency trained group; (a) Low, (b) Broad and (c) High. Nested plots show the 95% confidence intervals for the change for each measure of interest, namely (i) asymptote, (ii) midpoint and (iii) slope. Any results that are exclusively in the shaded are indicate a significant positive change (improvement). Results in the unshaded areas show a a significant negative change (worsening performance). Finally results that cross both zones (i.e include Zero) show no significant difference.

Table 4.3 Motion Coherence: Goodness of fit tests for each model. A mixed-effects model with the intercept as the only free parameter was compared with the model including the full psychometric function. -2 x the difference in log likelihood between the full and null model was calculated, and tested for significance against a chi-squared distribution, with degrees of freedom equal to the difference in the number of parameters for the two models.

Training Frequency	**Tested Frequency**	**-2LL**	**ΔDOF**	***p***
Low	Low	123.29	10	<0.0001
Low	Broad	146.17	10	<0.0001
Low	High	176.96	10	<0.0001
Broad	Low	161.96	10	<0.0001
Broad	Broad	158.93	10	<0.0001
Broad	High	168.91	10	<0.0001
High	Low	98.04	10	<0.0001
High	Broad	207.56	10	<0.0001
High	High	197.53	10	<0.0001

4.4.4 Transfer to global motion: Participant level variability

To investigate the lack of transfer to the trained task, a participant level plot (see Figure 4.7) illustrates the change in number of correct responses between pre- and post-training. This was calculated by subtracting the pre-assessment scores from the post-assessment scores. Changes above the dotted line at 0 show improved performance, while those below the line show worsened performance. The spread of scores for the broad/broad and broad/high condition show that one observer performed particularly poorly at higher coherence levels, but not persistently so. Therefore, the lack of overall improvement for the broad trained group is unlikely to be as a result of an outlier. However, the significantly worse performance in asymptotic performance may, to some degree, be as a result of this one individual.

4.4.5 Transfer to global motion: Summary

In summary, only the low frequency trained group provided robust evidence for transfer to trained and untrained motion conditions, with a moderate transfer to the broadband frequency condition.

4.4.6 Pre- and Post-Test Results for Contrast Sensitivity

We predicted that, should contrast sensitivity improve as a result of training on global motion it could only occur as a result of backward generalisation, as the two tasks are processed in cortically separate locations. We predicted transfer would be to limited conditions containing low frequencies, namely the broad and low trained groups.

a) Low-Frequency Training:

Results for the group trained with low-frequency stimuli are shown in Figure 4.8 (a). The model fits are summarised in Table 4.4. When tested with low-frequency stimuli, there was an increase in asymptote and a reduction in midpoint, but no change in slope. Overall, the results show an improvement in performance, particularly at high coherence levels. When tested with high-frequency stimuli, there was a reduction in asymptote, a small increase in midpoint, and no change in slope. These results reflect an overall worsening of performance.

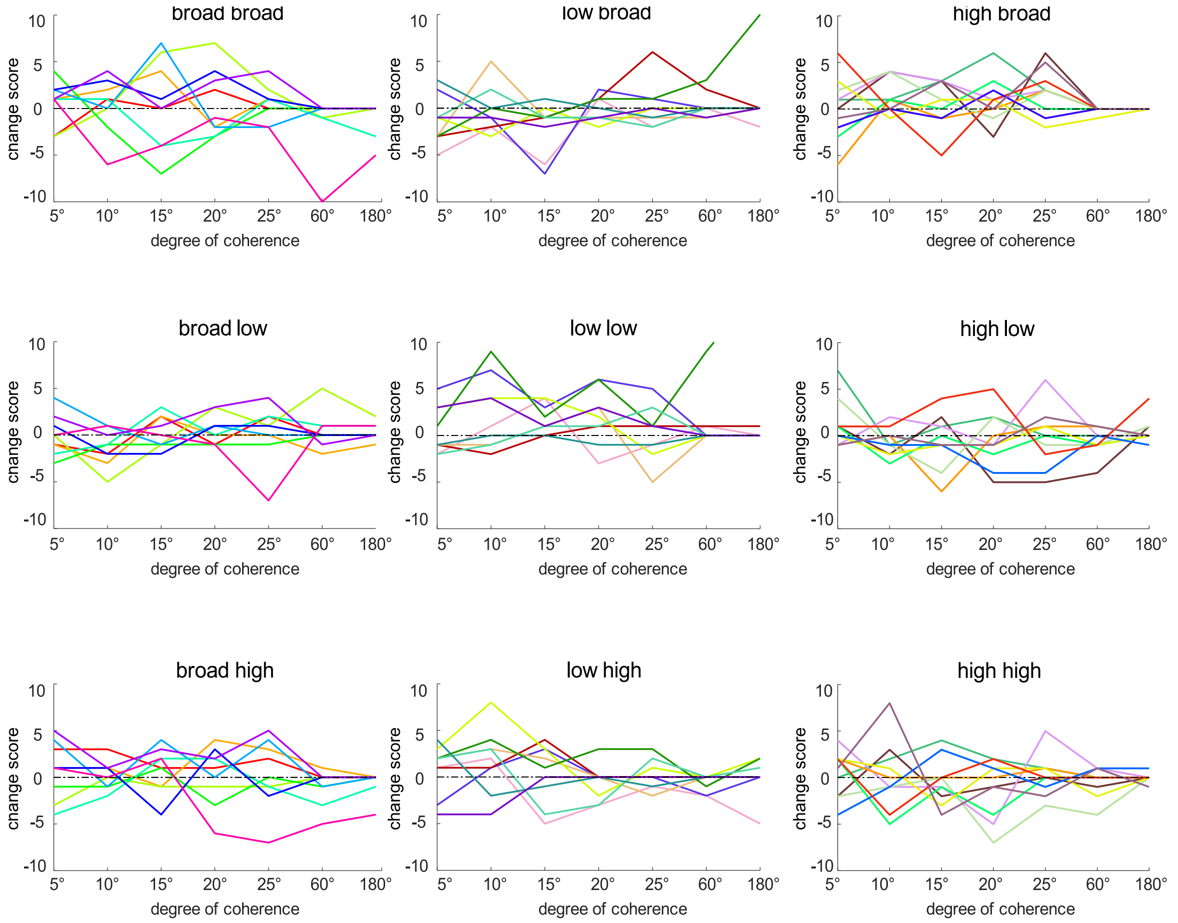


Figure. 4.7 Change scores (the difference between the number correct on the pre- and post-assessments) are plotted for each individual. Each observer is illustrated by a colour grouped by training and test condition (broad, low and high spatial frequency motion). Of particular interest is the broad trained group, which had a decline in performance post training. One observer performed worse at higher coherence levels after training, which explains the decline in asymptotic performance for the broad trained group. However, the overall spread of data suggest that on average post-assessment results did not differ significantly.

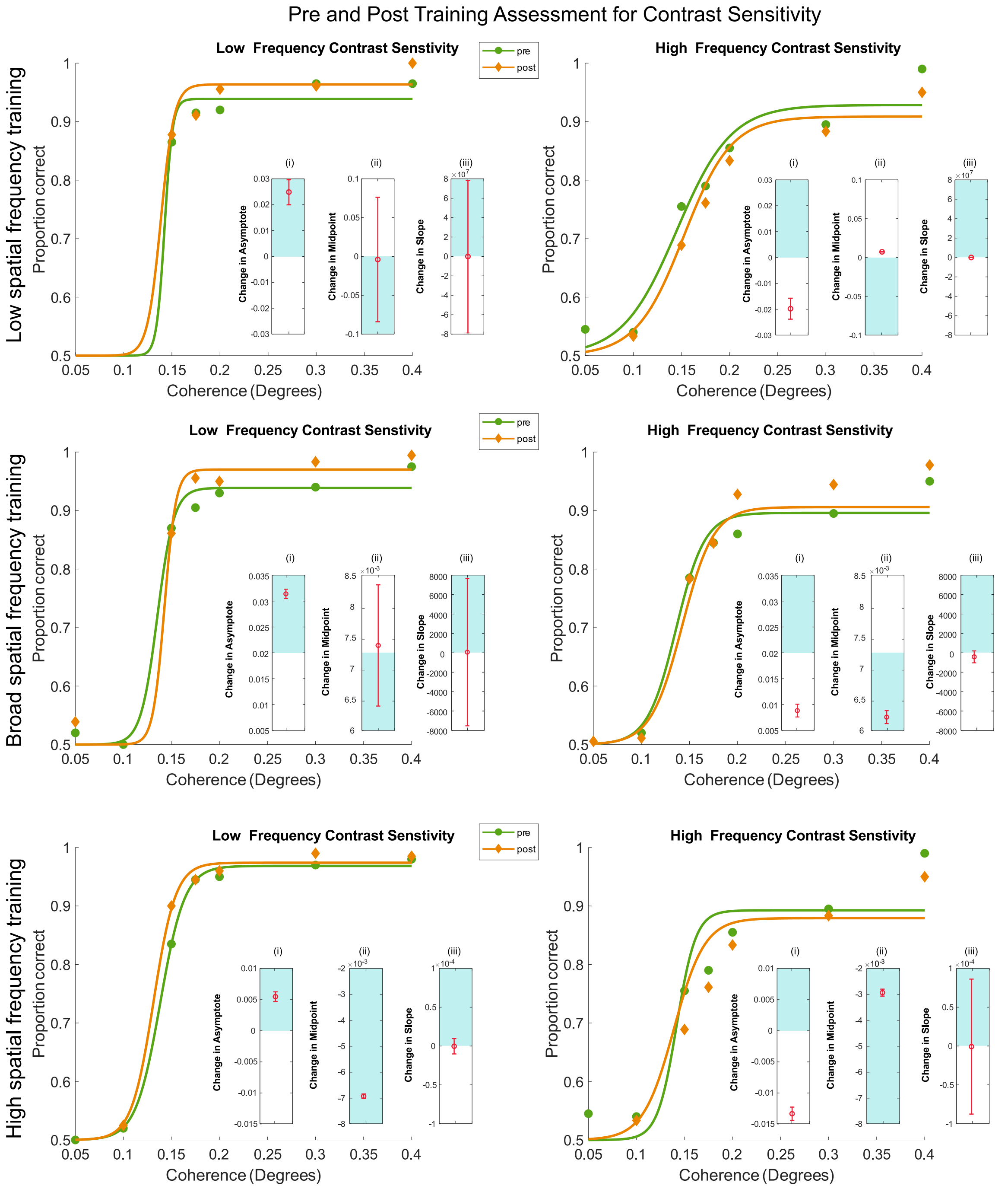


Figure. 4.8 Pre and Post Test Contrast Sensitivity (low and high spatial frequency) results following the five days training for each frequency trained group; (a) Low (a), (b) Broad and (c) High. Nested plots show the 95% confidence intervals for the change for each measure of interest, namely (i) asymptote, (ii) midpoint and (iii) slope. Any results that are exclusively in the shaded are indicate a significant positive change (improvement). Results in the unshaded areas show a significant negative change (worsening of performance). Finally results that cross both zones (i.e include Zero) show no significant difference.

b) Broad-Frequency Training:

Results for the group trained with broad-frequency stimuli are shown in Figure 4.8(b). The model fits are summarised in Table 4.4. When tested with low-frequency stimuli, there was an increase in asymptote, an increase in midpoint, but no change in slope. These results reflect an improvement in performance as a result of training, particularly at the higher contrast levels. When tested with high-frequency stimuli, there was an increase in asymptote, an increase in midpoint, and no change in slope. These results again reflect an improvement in performance as a result of training, particularly at the higher contrast levels.

c) High-Frequency Training:

Results for the group trained with high-frequency stimuli are shown in Figure 4.8(c). The model fits are summarised in Table 4.4. When tested with low-frequency stimuli, there was an increase in asymptote, a change in midpoint, but no change in slope. These results reflect a very modest improvement in performance following training. When tested with high-frequency stimuli, there was a reduction in asymptote, a reduction in midpoint, but no change in slope. These results represent a slight worsening of performance following training.

Table 4.4 Contrast Sensitivity : Goodness of fit tests for each model. A mixed-effects model with the intercept as the only free parameter was compared with the model including the full psychometric function. -2 x the difference in log likelihood between the full and null model was calculated, and tested for significance against a chi-squared distribution, with degrees of freedom equal to the difference in the number of parameters for the two models.

Training Frequency	**Tested Frequency**	**-2LL**	**ΔDOF**	***p***
Low	Low	234.87	8	<0.0001
Low	High	161.14	8	<0.0001
Broad	Low	268.60	8	<0.0001
Broad	High	206.31	8	<0.0001
High	Low	276.63	8	<0.0001
High	High	209.01	8	<0.0001

4.5 Discussion

4.5.1 Feedback and Perceptual Learning: Experiment 1

The purpose of the first experiment was to establish the necessity of performance feedback when undertaking a period of training on a task to discriminate the direction of global motion. Fahle and Edelman (1993) predicted that internal reinforcement could act as the teaching signal when performance feedback was absent and when the confidence is high (Jeter et al., 2009; Talluri et al., 2015). Therefore, perceptual learning should occur when training procedures include a mixture of easy and difficult trials. This was the case for the study conducted by Liu et al. (2012a) where learning occurred for easy and difficult trials without feedback. For this specific task, consistent with Seitz et al. (2006), the results of our study suggest that feedback was a requirement for learning. After ten days of training the group who received feedback improved significantly, while the performance for the no feedback group was ambiguous.

Our predictions were based on the findings of Liu et al. (2012a) who demonstrated that interleaving high accuracy (easy) trials and low accuracy (difficult) trials, resulted in perceptual learning without the need for feedback. Liu et al. (2012a) explain their findings using the Augmented Hebbian Re-weighting Model (AHRM) (Dosher & Lu, 2009). When external feedback is provided, the post-synaptic activation is shifted further in the correct direction, enforcing appropriate weight changes in the decision unit. However, when external feedback is absent the model uses the observer's internal response. In this situation learning is dependent on the level of difficulty of the task, and uses the observer's internal confidence to update the weights (Liu et al., 2012a). Where a task is easy, the weights still move, on average, in the correct direction (Liu et al., 2014). For a difficult task, the neural signal is weak and a clear indication of the appropriate changes that are required is absent. As a result the process of updating the decision weights is ineffective, and learning does not occur.

The results of our study showed that learning did not occur reliably without feedback, even when easy trials were presented. While our task spanned the 65-85% accuracy range (Liu et al., 2012a), there were some differences between the studies. Firstly, Liu et al. (2012a) used an adaptive staircase to track performance and stimuli were presented at either 85% or 65% accuracy, and at no other levels, throughout the experiment. Ours, on the other hand, did not include an adaptive staircase. Stimuli were randomly presented using MOCS. This method involves random stimulus selection from a predefined range of stimulus magnitudes (Kingdom & Prins, 2010). In our experiment observers were randomly presented 1 of 7 threshold levels. Interleaving these levels may influence how decision weights are updated and reduce the observer's confidence in their judgements, for example by disrupting observers' meta-cognitive judgements of perceptual confidence (Zylberberg et al., 2014) and thus their ability to selectively weight high-confidence trials in perceptual learning (Talluri et al., 2015).

Our findings bear some similarity to those found by Seitz et al. (2006), who also used MOCS. They found that learning did not occur without feedback, even when easy trials were presented. Observers trained on one of two tasks, either to discriminate the direction of low luminance motion stimuli, or to discriminate the orientation of a bar that had been masked in spatial noise. After training, both groups who had received feedback showed an improvement, while the groups without feedback did not improve. Seitz et al. (2006) note that many of the experiments investigating perceptual learning use adaptive staircase procedures, and even though easy trials were present, this difference may contribute to the difference in findings. Seitz et al. (2006) proposed that interleaving easy and difficult trials within the staircase may allow for “better bootstrapping" from easy to hard, compared to trials that are randomly presented.

A second difference between the study conducted by Liu et al. (2012a) and ours, was the type of task used. We used a global motion coherence task and not an orientation discrimination task (Liu et al., 2012a). These tasks are known to be processed at different levels of the visual hierarchy. Learning without feedback has been obtained for local motion tasks. Ball and Sekuler (1982) found that observers did not need feedback to improve on their direction discrimination task. In their study, observers were required to make a same/different judgement for two rapidly presented trials. In the “same” trials, motion took the same direction, and in the “different” trials the direction of motion varied by 3°. However, while the no feedback group did not receive trial-by-trial response feedback, they were rewarded with two cents for a correct response and had one cent deducted for each incorrect response, which may be construed as end of block feedback (Seitz et al., 2006), which has been found to be as effective as trial-by-trial feedback (Herzog & Fahle, 1997).

Learning without feedback has also been found for global motion coherence tasks (Vaina et al., 1998; Watanabe et al., 2001). However, none of these examples used the equivalent noise coherence task, but rather used the ratio of signal-to-noise coherence task. Levi et al. (2015), using an equivalent noise method of global motion, also found perceptual learning, however they used an adaptive staircase for training paired with trial-by-trial feedback. When investigating task-irrelevant learning, Watanabe et al. (2002) found that task-irrelevant local motion improved passively, but did not find the same for task-irrelevant global motion. They suggest that this is indicative of the lower levels of the visual hierarchy being more receptive to modification, when attention is limited. It was unexpected for the group trained without feedback that performance would deteriorate, and highlights interesting questions about motivation and performance over time as confidence and attention decline. The results of this study ultimately provide us with evidence that to obtain learning during training, using the equivalent noise motion coherence used by Levi et al. (2015), and presenting stimuli randomly using MOCS, our procedure should include trial-by-trial feedback during the training process. Furthermore, since learning did not occur without feedback, no feedback would be provided during the pre- and post-assessment phases.

4.5.2 Learning and Transfer of Global Motion: Experiment 2

Having established the necessity of feedback for learning, the objective of the main study was to investigate the specificity of spatial frequency tuning in perceptual learning for global motion. We trained three groups of observers on a global motion task with stimuli tuned to three different spatial frequency ranges (Broad, Low and High) and performed pre- and post-training assessments for all frequencies and high and low spatial frequency contrast detection.

4.5.3 Summary of Training

Our training results demonstrated that after the five day training period on a frequency specific global motion coherence task with trial-by-trial feedback, there was significant improvement in performance for all trained groups.

Pre- and Post-Assessment predictions

Based on the evidence for the frequency tuning of V5 (Amano et al., 2009; Bex & Dakin, 2002; Henriksson et al., 2008; Lui et al., 2007), we predicted that we might have obtained greater improvements (and transfer) between broad and low spatial frequencies, but limited or no transfer for the high spatial frequency trained group. Furthermore, we predicted that should transfer occur as a result of the reweighting of the cortically localised global motion mechanisms, then any transfer would reflect the broad, approximately low-pass spatial frequency tuning of global motion detectors. There would thus be most transfer when both training and test stimuli contained low spatial frequency components, and a moderate improvement for those trained or tested with broad frequency components. Since contrast discrimination is not processed in the same cortical location as global motion we predicted no transfer to contrast sensitivity would occur from any trained condition.

If transfer occurs as a result of backward generalisation, making use of the re-entrant (or feedback) connections as postulated by the Reverse Hierarchy Theory (Ahissar & Hochstein, 2004; Hochstein & Ahissar, 2002), we predicted transfer to show specificity to the frequency tuning of the global motion processing areas. Given the broadband low-pass frequency tuning of the motion processing areas, such as V5 and V3A, we predicted robust to moderate improvement where stimuli contain low spatial frequency components (learning to transfer from low to low; modest transfer of learning from low to broad, and broad to low and broad). In addition, we predicted frequency-specific improvement in contrast sensitivity with the broad trained group showing the most transfer, and improvements reflecting the spatial frequency tuning of the training stimuli.

Comment on the predictions

There is the possibility that transfer relies on forward connections from local to global motion mechanisms. However, given the tight frequency tuning at the local level and the well evidenced, highly specific characteristic of area V1, we would not make a prediction of transfer across frequencies occurring as a result of low level feed-forward mechanisms. However if learning did rely on the connections from local or global levels in a hierarchical fashion, we would expect frequency specific improvement on the matching spatial frequency motion condition (low to low, high to high, and broad to low and high). However, any transfer at the level of global motion, would depend on the global level mechanism so would again reflect the tuning properties of the global detectors.

Pre- and Post-Assessment findings

Our analysis of the pre- and post-assessment data found most support for transfer as a result of backwards generalisation, however not all our predictions were supported.

Firstly, while not an explicit prediction, we expected that the groups trained on their specific frequency conditions would naturally show improved performance for their trained frequency. However, when analysing the pre- and post-training motion results only the group trained using low spatial frequency Gabors displayed a significant learning, i.e. 'transfer to the trained task'. Notably, the broad trained group performed worse than they did at pre-assessment stage, and there was no change for the high trained group. Furthermore the only evidence of transfer to untrained conditions, was from the low frequency trained group on the broad test condition. Pre- and post-training assessment for contrast sensitivity found that there was a significant improvement exclusively for the low trained group on the low spatial frequency contrast condition. No further improvement was found for any other trained or tested frequency.

4.5.4 Evaluating the transfer from global motion

This study explored the effects of training on global motion and its transfer to other spatial frequencies and tasks. When assessing the post-training transfer to trained and untrained global motion frequencies, the *low spatial frequency trained group* was the only group to 'transfer to trained task' and transfer to another condition. Although the 75% threshold was worse for the broad test after training, the asymptotic performance was significantly better. Interestingly, for the high spatial frequency test, there was a significant improvement to the threshold although the slope was significantly shallower. This suggests that the sensitivity for lower coherence levels increased. Performance at the highest levels reached an asymptotic performance close to 1 at pre-assessment and remained unchanged at the post-assessment stage. Finally the shallower slope suggests a reduction in correct responses as a function of stimulus intensity.

The *high spatial frequency trained group* showed no improvement and no transfer to any other spatial frequency.

The most surprising results were those obtained from the *broad frequency trained group*. Asymptotic performance was significantly worse at the post test stage for their own trained frequency and the high frequency test, with no significant change in the low frequency test. There was no improvement in the slopes for any condition, and a small but significant improvement was found in 75% threshold for the high frequency test. Viewing individual performance revealed that the reduction in asymptotic performance may be accounted for by an outlier, however the outlier was unlikely to account for the overall absence of improvement in performance across all levels and measures, as there was no clear improvement evident for the other observers.

4.5.5 Comparisons to Levi et al.

Since Levi et al. (2015) reported improved contrast detection for low frequency drifting targets, we questioned if the improvement was specific to the temporal and spatial features of the training stimuli. Half the cells in layer 4Cα of V1, where improvement in contrast has been argued to occur (Sowden et al., 2002), are tuned for direction of motion (Hawken et al., 1988). It might be expected therefore that improvement in contrast sensitivity would be limited to moving stimuli.

Levi et al. (2015) evaluated three groups. The first trained on static, cardinally oriented Gabors, and made a horizontal or vertical direction judgement. The second trained on high intensity broadband frequency random dot coherence task, also making a direction judgement. The final group was untrained but undertook pre- and post-measures. We focus directly on the findings from the second training group which formed the basis for our initial enquiry.

Training

Levi et al. trained observers using a 3 up 1 down staircase, and provided auditory feedback on every trial. After 10 days training (300 trials daily), they reported an improvement in global motion perception. There was no specific pre- or post-measure for global motion integration (without trial-by-trial feedback). Improvement as a result of training was assessed by calculating the difference between performance on the first and last days of training. Our training was delivered using MOCS over 5 days with 420 trials (across 7 defined levels). To assess improvement across the training days we used a mixed-effects regression that includes all the responses over the 5 days (Moscatelli et al., 2012). Our training results are fully consistent with the findings of Levi et al.; training on a global motion integration task improves when training is paired with trial-by-trial feedback.

Pre and post-measures

As previously stated Levi et al. did not undertake pre- and post-measures for the trained task. Thus while our post-assessment for the broadband trained group showed that once feedback was removed the task was performed at prior levels, we are unable to make a direct comparison between their study and ours on this finding. Levi et al., tested four measures of transfer to contrast sensitivity: (i) orientation discrimination of static, horizontal or vertical non-flickering Gabors at varying spatial frequency (ii) orientation discrimination of moving horizontal or vertical gratings (temporal frequency of 10 hz) of varying spatial frequency (iii) direction discrimination (left/right) of moving gratings (temporal frequency of 10 hz) of varying spatial frequency and finally (iv) direction discrimination (left/right) of moving gratings (varying temporal frequency) and a spatial frequency held constant at 2 cycles per degree (medium frequency).

For (i) static ordinal Gabors, the group showed no improvement at any spatial frequency and the pre- and post-measures were no different from the untrained group. Only the moving gratings displayed an improvement in contrast sensitivity and these were predominantly found for low spatial frequency gratings.

The result from (i) is the most similar to our static contrast detection condition. Like Levi et al., our group trained on broadband global motion stimuli displayed no improvement for any tested spatial frequency for static contrast detection.

Since Levi et al. (2015) reported the biggest improvement for low spatial frequency stimuli exclusively in moving targets, this may suggest improvement was specific to the temporal and spatial features of the training stimuli. In contrast, our results suggest that when trained on low frequency global motion, robust learning and transfer occurred for stimuli that contained low frequency elements, including to a static contrast sensitivity task. Both of these findings are consistent with the low-pass frequency tuning of V5.

4.5.6 Contributions to understanding the mechanisms of perceptual learning

Specificity

We predicted that should global motion training improve contrast detection, given the frequency tuning of the early visual areas, we would expect specificity to the spatial frequency of training. Improvement was restricted to the low frequency contrast condition for the low trained group. The specificity of transfer to low frequencies is consistent with the low-pass frequency tuning of global motion areas such as V5 and V3/V3A influencing processing at V1 through feedback loops (Romei et al., 2016a). These results suggest however that there is some specificity for low frequency elements. This remains compatible with the view that global motion detectors pool for low frequency information. This may suggest that the re-entrant connections to V1, after training on global motion, only update low frequency channels.

Transfer

We predicted that if transfer occurred as a result of backward generalisation, then improvement would reflect the frequency tuning of the global motion detectors. The group trained on high spatial frequencies showed no improvement and no transfer. Since the representation of high spatial frequency content is attenuated in V5 (Henriksson et al., 2008), this may explain the lack of improvement in any post-assessment from the high or broad frequency trained groups. This may suggest any stimuli containing a high frequency element are attenuated for the purpose of the backward projections to V1 from global motion processing areas.

The broad trained group performed worse overall for almost all post-assessments. For the low spatial frequency contrast detection, there was evidence of a significantly steeper slope (a change in the rate of increase in proportion of correct responses with contrast), however this did not lead to a robust improvement in performance, and the average detection threshold was worse.

Transfer was only found for the low frequency trained group. This group showed improvement in low spatial frequency contrast detection, a significant increase in asymptotic performance for the broad frequency global motion test, and also improvement on their own trained task. This may suggest that after training on global motion only the low frequency channels are updated. The low frequency trained group will have experienced a high level of correlation in the activity in frequency channels between the feedforward and feedback connections. This joint activity may be important for perceptual learning, and is consistent with the proposal that there is an iterative interaction between the global motion processing areas and V1, since feedback from higher to lower areas engage the connections that mirror the features of the stimulus (Sillito et al., 2006).

4.5.7 Contribution to models of perceptual learning

The two main theories of visual perceptual learning suggest that learning either occurs through physiological changes to sensory neurons (Adini et al., 2002; Gilbert & Li, 2013; Gilbert et al., 2001), or as a reweighting (Dosher et al., 2013) of the decision units. Our findings are not able to distinguish between learning as a result of a physiological change or a higher-level computational reweighting. However neither model is able to account for the breadth of learning (and non-learning) and transfer within our results including the general absence of improvement and the occasional decline in performance at the post-assessment stage. The only difference between these components of the experiment was that trial-by-trial feedback was present for the training phase and absent for the testing phase. Our pilot study found that, for our methods, learning occurs with feedback, but not without. Additionally, Herzog and Fahle (1997) suggested that once feedback is removed, performance plateaus around the level last obtained with feedback, so our expectation was that observers would maintain the improvement achieved for their trained frequency. Neither model alone accounts for the ambiguous performance at the post-assessment stage or the decline in performance by the broad trained group. On the other hand, the specificity to low spatial frequencies is consistent with both models, given the frequency tuning of the global motion detectors.

While the two positions are pitched as competing theories, they may not be mutually exclusive. In an attempt to combine the two models, Solgi et al. (2013) model learning as a two-way process (descending and ascending). The model attributes training effects to the reweighting of connections between early and higher sensory areas, but crucially also as a result of increased neuronal activity representing the trained feature (Solgi et al., 2013). For transfer to occur, Solgi et al. (2013) argue there is an important role for the re-entrant connections from higher levels. Similarly, Bejjanki et al. (2011) propose a model that is computationally similar to the reweighting models, however it illustrates how changing the population codes in the early sensory areas can create similar changes in response to those made by high-level re-weighting models. Bejjanki et al. (2011) argue that their model captures the characteristics for perceptual learning for both behavioural and physiological changes. Sensory inputs improve the decision weights in the feed-forward connections, and improved probabilistic inference in early visual areas as result of the increased neural activity in the feedback network. This is compatible with our findings that suggest there the high level of consistency experienced for the group trained and tested on the low frequency conditions. Ultimately, a model of perceptual learning needs to account for multiple cognitive influences. Maniglia and Seitz (2018) argue that perceptual learning is not a detached or isolated process, instead multiple mechanisms react and interact to produce learning.

4.5.8 Afterthoughts, Speculation and Unresolved Points

We offer some speculative ideas that may help account for the general absence of improvement at the post-assessment stage. There is emerging evidence that interleaving random stimuli may disrupt perceptual learning(Aberg & Herzog, 2009). Yu et al. (2004) found that when interleaving random contrasted Gabor stimuli, such as when using MOCS, learning did not occur, but did when only one contrast was presented at a time. Yu et al. (2004) termed the interleaving of contrast stimuli as “contrast roving”, and suggest that the effect of non-learning as a result of contrast roving implies that a temporally organised pattern of stimulus presentation may be required for perceptual learning to occur. Additionally, Kuai et al. (2005) found that temporal patterning was evident for motion direction discrimination. Observers were either trained using randomly varied trials or with a fixed temporal pattern of exposure (changing in a sequential clockwise direction). The first group showed no improvement in discrimination, however the second group with the fixed temporal exposure did. Cong and Zhang (2014) showed that a semantic tag associated with each stimulus presentation enabled significant learning. Sequential tags such as A,B,C,D and 1,2,3,4 which contain semantic and sequential identity information both resulted in significant learning. Cong and Zhang (2014) suggest that tagging and temporal sequencing may help direct the system to switch attention to the neuronal set responsible to the specific stimulus. These are consistent with the experiments conducted that also found learning did not occur without feedback using MOCS (Seitz et al., 2006).

In addition, there is also evidence that interleaving easy and difficult trials during perceptual training may only lead to a temporary improvement, that (a) does not persist over time and (b) disappears when easy trials are removed (Lin & Murray, 2014, 2015). In a difficult task without feedback learning does not occur, however it does with feedback (Liu et al., 2012a). Removing the feedback may be equivalent to removing the easy trials, and therefore reducing confidence, which is known to be an important factor in learning and perception (Asher & Hibbard, 2018). Lin et al. (2017) presented observers with a single block of mixed easy and difficult trials which led to rapid improvement within the block. However, after removing the easy trials and presenting only the difficult trials performance slowly returned to prior levels. This is consistent with findings that internal reinforcement could act as the teaching signal when confidence is high (Jeter et al., 2009; Talluri et al., 2015).

In our study, improvement in the trained task was sustained only while feedback was present. At the post-assessment stage, the combination of a difficult task, the interleaved coherence levels and the general low decision confidence in the absence of feedback may have inhibited 'transfer to the trained task' and, like Lin and Dosher (2017), performance returned to baseline. It is thus a possibility that the amount of learning for the broad and high spatial frequency trained groups was insufficient to overcome the reduction in performance when feedback was removed.

An exception to this was the improvement in performance and 'transfer to the trained task', for the low frequency trained group. Since learning and transfer to trained task were evident for the low spatial frequency trained group, this may suggest an interaction between the amount of learning, and the benefit of feedback. Furthermore, the improvement by the low frequency group over the five training days shows the biggest move from baseline performance, and thus potentially benefited from more learning, which was robust, persistent and thus incurred less disruption by the absence of feedback.

4.6 Conclusions

Our results support the ideas from the Reverse Hierarchy Theory that feedback connections from higher levels back to early visual processing areas may be involved in facilitating learning that transfers (Ahissar & Hochstein, 2004; Hochstein & Ahissar, 2002), and are consistent with models that suggest that the re-entrant connections play a vital role in learning and transfer (Bejjanki et al., 2011; Solgi et al., 2013).

Our first experiment suggested that for perceptual learning of an equivalent noise global motion task, presented using MOCS, feedback was a necessity. However after training with feedback, there was little evidence of a robust improvement when feedback was removed at the post-assessment stage. The only condition that showed unambiguous learning and transfer was the low spatial frequency trained group. Our findings are in line with current research which finds global motion detectors pool for low spatial frequencies.

Our main experiment found robust learning and transfer for the group trained with low spatial frequency specific global motion gabors. This provided evidence of frequency specific transfer from a higher-level global motion task to a low-level static untrained local task. The frequency tuning of these results is consistent with perceptual learning which depends on the global stage of motion processing.

Finally, the frequency tuning of these results suggests that the mechanism of transfer of global motion is dependent on backward generalisation, and specific to the frequency tuning of the global motion detectors. Transfer exhibited a high specificity towards low spatial frequencies, consistent with the attenuation found in the processing of global motion.

Our study supports the idea that the perception of global motion shows a bias for low spatial frequencies and suggests the backward projections and the re-entrant connections from global motion processing areas to V1 play a crucial role in transfer to trained and untrained tasks.

4.7 Acknowledgments

This research was funded by a grants from ESSEXLab and PsyPAG to J.A.

Chapter 5

Robust learning without feedback, with interleaved easy and difficult trials, for local and global motion and form

5.1 Introduction

The necessity for external reinforcement during perceptual training is a topic of rich theoretical and empirical debate (Dosher & Lu, 1999; Herzog et al., 2006; Herzog & Fahle, 1997; Seitz et al., 2006; Shibata et al., 2009). External reinforcement, such as response feedback, occurs when the learning environment provides an observer with an indication of the quality of their behaviour or response (Dosher & Lu, 2017). In the laboratory, this is often provided with auditory tones, to indicate whether the response just given was correct or incorrect. There are many studies that find that providing an observer with external feedback increases the rate of perceptual learning (Garcia et al., 2013; Herzog & Fahle, 1997). Training paired with feedback increases performance and the efficiency (Dobres & Watanabe, 2012; Herzog & Fahle, 1998) of learning. Feedback is often presented at the end of each trial, but it can also be provided at the end of a training block (Herzog & Fahle, 1997; Shibata et al., 2009). While trial-by-trial feedback is more effective than end of block feedback, the latter has still been found to improve the rate of learning (Shibata et al., 2009). Learning has also been found to occur in the absence of external feedback (Ball & Sekuler, 1982; Fahle, 2005; Liu et al., 2012a; Petrov et al., 2006a; Shiu & Pashler, 1992; Vaina et al., 1998), specifically when the task is easy, but not when it is difficult (Liu et al., 2010). Evidence suggests that learning without feedback sometimes occurs when easy and difficult trials are interleaved (Fahle & Edelman, 1993; Liu et al., 2012a), but not always (Asher et al., 2018b; Seitz et al., 2006). However, human and non-human animals develop their perceptual skills by actively exploring the world around them without the need for constant feedback. Feedback is not routinely the primary interest in perceptual learning studies, and there have until very recently only been a handful of studies that explicitly manipulate the use feedback when learning (Asher et al., 2018b; Liu et al., 2012a; Seitz et al., 2006). Furthermore, these studies have produced some conflicting and equivocal results. The studies that do manipulate feedback are difficult to compare, owing to the mixtures of stimuli, levels of processing, task difficulty, and performance confidence and accuracy (Liu et al., 2012a). For example, some studies may track performance over time using varying stimulus intensities (Petrov et al., 2005, 2006a) and measure how performance varies at each level. Others may keep the stimulus level constant and measure how performance changes over time (Herzog & Fahle, 1997;

Vaina et al., 1995), or measure the change in threshold while holding stimulus intensity fixed at a set level of performance accuracy (Ahissar & Hochstein, 1997; Karni & Sagi, 1991; Liu et al., 2012a).

5.1.1 Supervised and unsupervised learning

Perceptual learning in human and non-human animals is often modelled using computational neural network models to test and evaluate the complex interactions between learning and feedback (Herzog & Fahle, 1998; Petrov et al., 2006b; Tsodyks & Gilbert, 2004). These models use learning rules to calculate a change in weights between a higher level decision unit its low-level sensory representation (Herzog & Fahle, 1998; Petrov et al., 2006b). Learning rules in machine learning are supervised or unsupervised; supervised learning requires a teaching signal, while unsupervised learning does not (Fahle, 2002). Similarly, perceptual learning is likely to occur as a result of exploiting the correlations and neuronal connections via Hebbian learning (Gilbert et al., 2001; Kempter et al., 1999). Like a well trodden path, neural circuits in the brain rely on repeatedly correlated activity in neural firing (Song et al., 2000). The correlations in neural firing rates between pre- and post-synaptic neurons drive the development of the receptive fields for neurons (Kempter et al., 1999). Learning rules correlate the behaviour between pre- and post-synaptic firing and update the connections between the visual and decision making areas. The weight of a synapse is characterised by the strength of its connection to another synapse, and the relative timing of neural firing from one neuron to the next (Song et al., 2000).

Supervised networks operate on a feedforward system that requires explicit external feedback (Petrov et al., 2005). Neurons in feedforward networks are organised in hierarchies, where information flows in one direction, from one level to another (Tsodyks & Gilbert, 2004). During stimulus presentation, the learning rule adjusts the weights used to classify the stimulus. A discrepancy between the layers usually triggers a change, which is provided during the learning process with a teaching signal which strengthens the connections between these layers and updates the neural connection weights, thus reducing the discrepancy between the actual and target activation (Petrov et al., 2005; Tsodyks & Gilbert, 2004). In supervised learning, the teaching signal strengthens the connections for that decision weight shifting it in the correct direction. Thus, the aim of the teaching signal is to minimise the error produced after the presentation of the stimulus (Herzog & Fahle, 1999).

In unsupervised networks, information is transmitted within and between higher and lower levels, increasing connection weights on the basis of co-activation (Tsodyks & Gilbert, 2004). They are described through Hebbian Learning rules (Herzog & Fahle, 1997). Therefore, learning occurs as a result of exposure to a stimulus, which strengthens neural connections (Herzog & Fahle, 1997). Since learning within this model is stimulus driven, occurring without top down influence, learning occurs independently from feedback signals (Herzog & Fahle, 1997; Petrov et al., 2005; Tsodyks et al., 2004). In this type of network, learning is considered unsupervised. This type of learning can, for example, depend on correlations in the responses of sensory neurons and decision units, in which case learning may operate in an unsupervised manner, through mere exposure to appropriate stimuli (Herzog & Fahle, 1997; Seitz et al., 2006; Shiu & Pashler, 1992). However, not all studies have obtained evidence of learning from exposure alone (Garcia et al., 2013; Herzog & Fahle, 1997; Seitz et al., 2006; Shiu & Pashler, 1992).

Teaching signals may also be obtained from internal reinforcement (Herzog & Fahle, 1997; Petrov et al., 2005), whereby easy stimuli act as a teaching signal (Fahle & Edelman, 1993). Easy to discriminate stimuli act as an exemplar for more difficult stimuli (Herzog & Fahle, 1997), and an internal signal of confidence provides the ‘bootstrapping process’ to aide in the detection of more difficult stimuli (Fahle & Edelman, 1993). In the context of Hebbian learning, an observer’s internal conviction would update the decision weights (Liu et al., 2014), and a high correlation would exist between the decision weight and the decision made when confidence is high. Learning is thus dependent on the level of difficulty of the task, or sufficient exemplars (Ahissar & Hochstein, 1997). However, the ambiguous and contradictory psychophysical findings from human observers suggest that learning is neither exclusively supervised nor unsupervised, and more likely a combination of both (Herzog & Fahle, 1998).

Sutton and Barto (p4 2011) argued that supervised learning is by definition learning obtained via an ‘expert supervisor’. Others have proposed that human learning is neither supervised nor unsupervised, but modelled by reinforcement learning (Law & Gold, 2009; Sutton & Barto, 2011). Reinforcement learning occurs as a process of trial and error while interacting with the environment (Sutton & Barto, 2011). Whether human learning with internal feedback can be considered a teaching signal is unclear. However, where a task is very difficult and does not contain sufficient numbers of easy exemplars, learning does not occur (Liu et al., 2012a). Within a neural network model, this is explained by random or slow updating of the decision weights. Thus, where response feedback is absent, learning is dependent on the level of difficulty of the task (Ahissar & Hochstein, 1997).

5.1.2 Easy and difficult trials

A wealth of perceptual learning studies support the idea that training on easy to discriminate trials facilitates learning for difficult trials (Ahissar & Hochstein, 1997; Garcia et al., 2013; Jeter et al., 2009; Liu et al., 2010, 2012a; Liu & Weinshall, 2000). In an orientation discrimination task, Ahissar and Hochstein (2004) manipulated task difficultly by decreasing the stimulus processing time available between the presentation of a target and a distractor, and found that, as difficulty increased, improvement was slower and specific to the trained orientation. However, in the conditions with increased time between stimuli and distractors, learning was quicker and transferred across orientations. Training was also found to be more effective when there was a gradual transition from easy to difficult trials (Ahissar & Hochstein, 2004). Investigating the time course for detecting global form in Glass patterns, (Garcia et al., 2013) found that, consistent with prior evidence, training on easy tasks facilitated performance on later difficult trials. However, the results also suggested that performance improvements were greater with those participants trained on difficult tasks, even when those trained with easy tasks received twice as much training (Garcia et al., 2013). Similarly, Liu et al. (2012a) found that when training an orientation discrimination task that was set at a low accuracy (65% correct), learning occurred when external feedback was provided, but not when feedback was absent. In contrast, when high accuracy (85% correct) trials were used, groups improved whether feedback was present or not. However, in contrast to a gradual transition from easy to difficult (Ahissar & Hochstein, 2004; Garcia et al., 2013), in this study high and low accuracy trials were mixed, and performance improved for both high and low accuracy trials, for both the feedback and the no feedback conditions(Liu et al., 2012a).

5.1.3 Levels of processing

At the lowest levels of the visual processing hierarchy, V1 and V2 represent simple dimensions such as position, orientation and direction of motion, and respond only to a small area of the visual field (Gilbert et al., 2001). On the other hand, receptive fields in higher cortical areas are larger, and represent more complex features that are pooled over space, time, speed, direction and frequency (Burr & Thompson, 2011; Felleman & Van Essen, 1987; Furlan & Smith, 2016; Gilbert et al., 2001; Hubel & Wiesel, 1965; Mikami et al., 1986; Movshon et al., 1978; Sillito et al., 2006; Zeki, 1974).

Since the receptive fields in lower levels of the visual cortex only represent a *localised* area of the visual field (Born & Bradley, 2005; Gur & Snodderly, 2007; Hu et al., 2018; Wandell & Smirnakis, 2009), they are predominately responsible for processing tasks where the stimulus attributes match their specific tuning. These are often referred to as *local tasks* in psychophysical studies (Edwards et al., 2013; Nishida, 2011). Local tasks do not necessarily include external noise, and for example may be comprised simply of a signal to be detected or discriminated, in various levels of intensity or difficulty (for example see figure Figure 5.1 (a - b)). This illustrates a local motion task where an observer is required to make a left/right direction of motion decision relative to the vertical midpoint. In this schematic, all the dots are moving leftwards. Whether you are presented with 20 dots or 1 dot, the direction of motion can be discriminated. In this instance, difficultly can be increased by reducing the angle of tilt from the vertical mid-point at 0°. Accuracy in detecting the direction increases with (i) more dots providing more signal (ii) a larger tilt from the vertical mid-point.

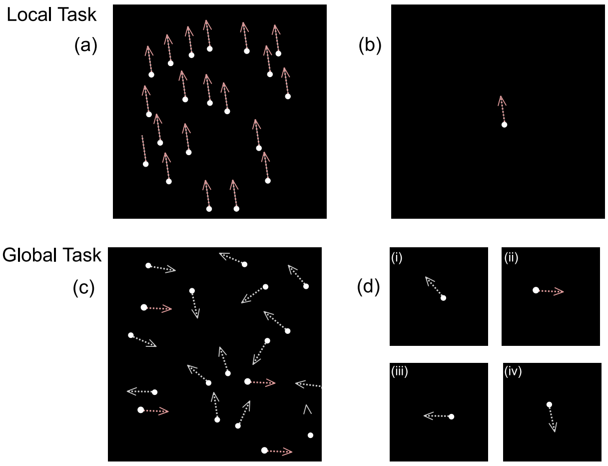


Figure. 5.1 A local motion task (a) that requires an observer to make a left right discrimination decision from the vertical mid-point. All the dots move coherently upwards and slightly towards the left, however in image (b) only supplying one dot does not impede direction discrimination. In contrast when using a global motion task (c), for an observer to make a left right decision, they need to integrate the signal and segment it from the noise. In this case removing all the dots, apart from one, makes this task impossible to complete. In (d i-iv), irrelevant of whether a signal dot or noise dot is presented a correct response would be no better than chance

In contrast, the receptive fields of neurons in higher cortical areas such as V4 and MT pool over stimulus properties across a broad area of the visual field, in order to provide a *holistic* view (Amano et al., 2009; Bex & Dakin, 2002; Braddick, 1993; Burr & Thompson, 2011; Gilbert et al., 2001; Nishida, 2011). Unlike the clear signal provided in local tasks, a general direction of motion cannot be determined by the trajectory of a single dot (Tibber et al., 2014). Tasks that can only be resolved through integration and segregation of coherent or conflicting information are often referred to as *global tasks* (Burr & Thompson, 2011; Garcia et al., 2013; Nankoo et al., 2015; Nishida, 2011). Global tasks such as motion or form coherence are investigated using stimuli with conflicting signals. To process complex motion, the signals from the first-stage need to be integrated over a wider spatial area to provide a global representation of motion (Adelson & Movshon, 1982; Wilson et al., 1992). Figure 5.1 (c) illustrates a global motion task where an observer is required to make a left/right direction of motion judgement based on the pooled motion. In this task a percentage of signal dots would move coherently either leftwards or rightwards, and distractors (noise) would move around randomly. Performance increases as the proportion of signal dots increases, thus a more difficult task would have more noise dots. As depicted in panels Figure 5.1 d (i-iv) only presenting a single dot, whether signal or noise, will make identifying the direction of motion from that trial a matter of chance.

5.1.4 Demands of different tasks and processing mechanisms

Little is understood about the mechanisms driving learning at the various levels of hierarchy within the visual system (Law & Gold, 2008). The visual processing hierarchy is structurally connected by neural pathways that are either feedforward or feedback connections. These pathways borrow terms from computational models whereby feedforward connections send stimulus driven sensory information *up* the visual hierarchy from the lowest cortical areas (bottom) to the higher level areas (Markov et al., 2014). Thus *bottom-up* perception can be thought of as representing the feature specific content of the stimulus. For example, when viewing for the first time, the image in Figure 5.2 appears comprised exclusively of randomly placed black splodges on a white canvas. However, perception is influenced by stimulus driven processing *bottom-up* and *top-down* processing (Gilbert & Li, 2013).

Feedback connections (re-entrant) are those pathways from the higher levels of the visual hierarchy *down* to the lowest levels (Markov et al., 2014). *Top-down* influences such as attention (Gilbert & Li, 2013; Westheimer et al., 2000), prior knowledge (Barlow, 1990; Zhaoping, 2017) and confidence (Asher & Hibbard, 2018; Talluri et al., 2015) can influence perception. A practical example of this is the Eureka or Aha! Effect, a second look at Figure 5.2 reveals an embedded figure, which is more rapidly identified once the viewer has been primed (Ahissar & Hochstein, 2004; Gilbert & Sigman, 2007) to see a dog. This process is often permanent, even after a single very brief exposure to a prime (Ahissar & Hochstein, 2004). In terms of of the bottom-up sensory responses, the neural activity would remain identical at both low and high level visual areas. What changes is a top-down, higher level representation of the scene.

Figure. 5.2 The Eureka effect explains how priming, or providing a hint, can help first time viewers make sense of an ambiguous scene. The black and white splodges are instantly identifiable when advised to look for the embedded figure of a spotted Dalmatian smelling or drinking near a tree. [Figure obtained licence free via Wikimedia Commons]

It has been proposed that visual perceptual learning may involve two theoretical stages (Shibata et al., 2014), whereby low-level (local stage) learning is stimulus driven, and higher-level (global stages) are dependent on the demands of the task, and other top-down influences such as attention. Stimulus driven learning (learning from exposure) is often evaluated by using task relevant and task irrelevant paradigms. Task-irrelevant contents are those features that do not convey information that is required to complete the task (Seitz & Watanabe, 2009). While learning from exposure alone is not sufficient to give rise to improved performance (Garcia et al., 2013; Herzog & Fahle, 1997; Seitz et al., 2006; Shiu & Pashler, 1992), using training tasks that have an embedded irrelevant feature is argued to trigger improvements in the readouts for both features (Solgi et al., 2013). These paradigms have found that features irrelevant to the task can be learned as a result of persistent exposure (Seitz et al., 2005; Seitz & Watanabe, 2005, 2009; Watanabe et al., 2002, 2001). In a study using dynamic random dot motion, where the stimulus contained both a local and a global motion element (Watanabe et al., 2002), task irrelevant improvement only occurred for low-level local stimuli. The stimulus was designed so that the direction of local motion varied, and the direction of global motion remained fixed, but never overlapped (Seitz & Watanabe, 2009; Watanabe et al., 2002). When global motion was the task-relevant goal of the dynamic motion training stimulus, performance improved as expected. On the other hand, when the dynamic motion task was presented as a task-irrelevant stimulus (while reporting letter sequences), performance improved for local but not global motion (Watanabe et al., 2002). They suggest that this is indicative of the lower, local level of the visual hierarchy being more receptive to stimulus driven *bottom-up* modification. In contrast, when the task was more demanding the irrelevant features (the global motion element), which are influenced by *top-down* processing, are ignored (Ahissar & Hochstein, 2004; Shibata et al., 2014).

Perceptual learning has been found to occur when participants are provided with fake feedback (Shibata et al., 2009). Participants made a same/different decision, and the first and second interval were separated by an inter-interval gap of between 500-1000ms. Participants were provided with modulated fake feedback based on the gradient of responses (see Shibata et al. (2009) for the full methods). The groups trained with accurate feedback or no feedback showed significant learning, however the group with the modulated feedback demonstrated an overall larger improvement in performance. This also suggests that learning can be manipulated when uncertainty is high. Furthermore, uncertainty may be even greater for an ambiguous global level task, and external feedback may be weighted as more reliable than ones internal feedback. Perhaps then, the requirement for successful perceptual learning differs throughout the hierarchy where global tasks may depend more on the presence of feedback for learning.

When assessing the time course of learning for global form (Garcia et al., 2013), difficult and easy stimuli were defined by adjusting the spiral angle of radial and concentric Glass patterns. These were not interleaved, and each group was trained on either difficult or easy to detect Glass patterns, and then tested on a difficult Glass pattern task. For groups trained on difficult tasks, the time course for learning suggested that their performance increased within each session, but plateaued by the final day of training. For the group trained only with easy stimuli, within session learning was negligible however performance increased from one session to the next. Finally, the group trained with easy-to-detect stimuli showed a lower overall improvement (Garcia et al., 2013). This suggests that training (for 3 days) to discriminate more difficult global form patterns resulted in stronger behavioural improvements than training with easy trials. This was not consistent with our study of global motion (Asher et al., 2018b) where post-training assessments suggested that there were no behavioural improvements after five days of training without feedback. This may suggest differences in learning mechanisms that process form and motion. Based on the findings reported by Garcia et al. (2013) learning may occur more quickly for form than motion.

5.1.5 Methods of presentation

A difference between the studies that did (Liu et al., 2012a) and did not (Asher et al., 2018b; Seitz et al., 2006) find learning with interleaved difficulties of trials was the method of presentation, with the former using an adaptive staircase and both the latter studies using MOCS (these methods are discussed in detail in Chapter 2). MOCS involves repeatedly presenting randomly interleaved stimuli that are drawn from a predefined set of levels (Kingdom & Prins, 2010). Staircases use a customisable, rule-based algorithm to select the level of stimulus that is presented on each trial. Dependent on the type of staircase used, the staircase can converge on a desired threshold (see Levitt (1971) for more detail). For example if the desired threshold is 75% correct (mid-way between chance and 100% correct in a 2AFC, (see Table 2.1 in Chapter 2) the staircase creates a stepped presentation from easy to difficult to detect stimuli, in order to converge on the stimulus intensity required in order to obtain 75% accuracy.

It has been argued that because many staircases present a graduating increase in difficulty, detecting easy stimuli facilitates the detection of difficult stimuli, such that the sequence of presentation may provide bootstrapping between the levels of difficulty in stimulus intensity (Seitz et al., 2006). If this is the case, then staircasing may be comparable to the case of easy trials providing a support mechanism for learning of difficult trials. However, little is known about the exact proportion of easy and difficult trials, or how to quantify ‘easy’ and ‘difficult’, in order to best facilitate learning. Liu et al. (2012a) trained observers using a staircase to ensure that training was maintained at a performance accuracy of either 65% and 85% for difficult and easy trials respectively, or both, when the two levels were interleaved. Using the same range of levels of accuracy (between 65% and 85%), our pilot study (Asher et al., 2018b) presented a range of intensities between these two points. Furthermore, in the second experiment post- training assessment tests found that even after five days of training with interleaved easy and difficult trials paired with trial-by-trial feedback, performance returned to the pre-training baseline thresholds in all but one condition.

5.1.6 Aims and predictions

Based on the findings that mixing easy and difficult trials facilitates learning without feedback and the ambiguity between successful (Liu et al., 2012a) and unsuccessful (Asher et al., 2018b; Seitz et al., 2006) learning, the purpose of this study was to explore the need for external feedback in perceptual training for local and global form and motion tasks. In order to disentangle the conflicting results, we considered a number of factors that may produce different results including (i) the level of processing and the type of task (or task demand),

(ii) the provision of trial-by-trial feedback and (iii) the method of presentation of the stimuli (e.g. adaptive staircase or MOCS). Based on previous results and theoretical considerations, we predicted the following:

1. Given the nature of stimulus driven improvements, local tasks should improve in the absence feedback.

2. If there is a difference in higher level processing between form and motion, we should expect to see some difference in the global no feedback conditions between the two tasks. Specifically, we might expect improvement is faster for the form conditions.

3. If the areas responsible for integration and segmentation require external feedback for learning, then the groups trained without feedback should show no (or less) improvement.

4. Finally, if the stimulus presentation type influences learning we may expect a difference in learning without feedback between the MOCS and staircase conditions. Specifically, we expect learning without external feedback to be more effective when a staircase technique is used to present stimuli at distinct easy and difficult levels of performance, rather than when using MOCS.

5.2 General Methods

5.2.1 Participants

140 observers from the University of Essex participated in the study, in one of ten conditions, resulting in a total of 14 participants in each condition. All observers had normal or corrected-to-normal vision and were paid for their participation or received course credit. Procedures were approved by the University of Essex University Ethics Committee, and all participants gave informed written consent.

5.2.2 Apparatus

Stimuli were presented on a 27" iMac with a display resolution of 2560x1440 pixels and 60 Hz refresh rate. Stimuli were generated and presented with Matlab 2015a, using the Psychophysics Toolbox extensions (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997) on a 2.7 Ghz IMac running OSX 10.9.5. Stimuli were viewed from a distance of 500 mm, and one pixel subtended 1.6’(arcminutes).

5.2.3 General Stimuli and Procedural information

Stimuli were 4.8’ $\times$ 4.8’ square white dots displayed within a black 7.95 $^\circ\times$ 7.95° square region in the centre of the black screen. For the motion conditions, there were 255 white dots while for the form tasks there were 300 white dot pairs (600 dots). Half of the observers were provided with trial-by-trial feedback during the training blocks. When feedback was present, this was provided as an immediate auditory beep after each trial, a high pitched tone for a correct response (2000Hz for 100ms), and a low pitched tone for an incorrect response (200 Hz for 400 ms). No feedback was provided for any group during the pre- and post-assessment phases.

The general procedure for all conditions was identical. Participants were trained for three consecutive days on one condition only, either with or without feedback. Each participant completed daily baseline measures before and after training for the same task, without feedback. All participants were briefed orally, read written instructions on the task and finally completed a short demonstration. All testing was performed in a dark room and participants were regularly prompted to check their distance from the screen with a measured piece of string that was affixed to the monitor. A 30 second break was provided before and after the training block, and 10 second breaks were provided in between each of the 12 training blocks. While breaks were timed, participants were advised they may take longer should they need to.

Experiment 1 - Local and Global Form

The individual dots in each pair were separated by 9.6’ (see Figure 5.3 (a)). Form was defined by manipulating the direction from one dot to the other. So as not to have overlap, dots were rendered on an invisible grid at the centre of the screen. Each of the square regions was $0.53 \times 0.53\,^{\circ}$ and contained a single randomly positioned dot pair. Each stimulus was presented for 1 second and the next trial did not begin until a response was made. A fixation cross would appear in the centre of the screen for 500 ms, then the next trial after a further 500ms pause.

The pre- and post-assessment for all conditions consisted of a left or right Two Alternative Forced Choice (2AFC) decision. Stimuli were presented using the method of constant stimuli (MOCS), and participants responded pressing the left or right arrows on the keyboard. The order of trials was randomised, and contained 30 repetitions for each level (270 trials). Testing was performed before and after training each day. For the local task, dots were either oriented on the vertical 0°, or tilted at one of the following angles ± 5°, 2.5°, 1.25°, 0.625°. Participants were required to indicate whether dots were oriented towards the left or right of the vertical mid-point. For the global task dot pairs were randomly designated as signal or noise. The number of signal dots for each trial was randomly selected from the predefined levels (5, 10, 15, 37, 60, 75, 100, 125, 150). The remaining dots were then designated as noise dots for that trial. All signal dots were orientated either to left or right of vertical at ± 20° for each trial. Each noise dot pair was randomly tilted between 0-360°. Example stimuli are shown in Figure 5.3 (b). Participants were required to identify the direction of tilt (at either ± 20°) provided by the oriented signal dots.

A cumulative Gaussian curve was fit to participants' responses from their pre-assessment data each day using the Palamedes toolbox (Prins & Kingdom, 2009) in order to determine 65% and 85% correct discrimination thresholds for difficult and easy trials respectively (as used by Liu et al. (2012a)) for their training blocks.

The training task used the same stimuli as the pre- and post conditions, however stimuli were presented at one of two levels either at 65% and 85% accuracy, as determined at the levels described above. A QUEST procedure was used to maintain the stimuli at the appropriate performance level throughout the block (Watson & Pelli, 1983). Trials from the two staircases were interleaved. At the start of each new block, thresholds returned to the original level. Training consisted of 12 blocks with 40 trials for each level, with a total of 80 trials within each block (960 trials per day).

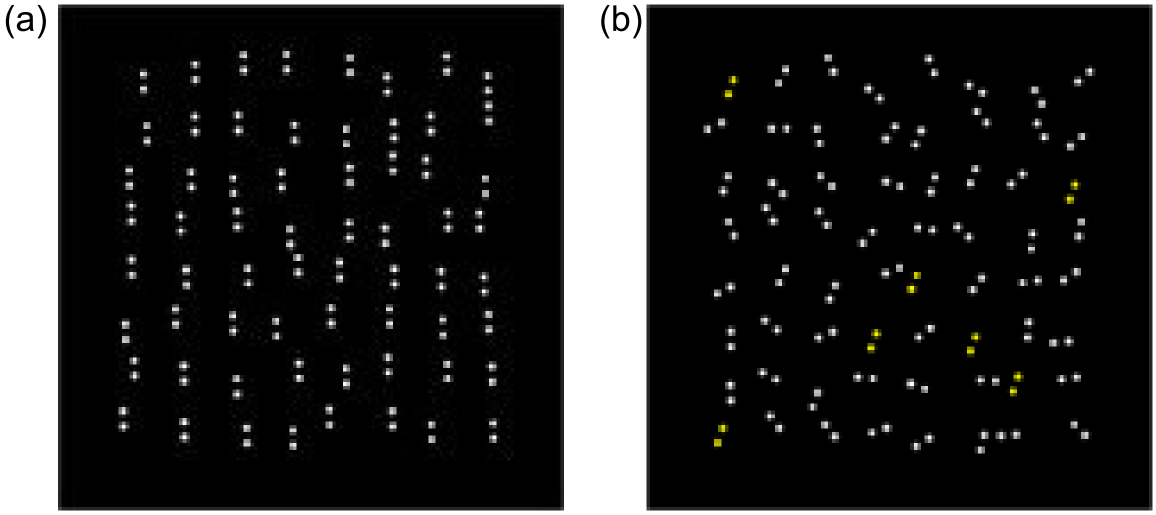


Figure. 5.3 Left: Local Form stimuli where all dots are oriented leftwards at 1.25° from the vertical. Right: Global form stimuli where 7 (indicated in yellow) are signal dot pairs (oriented 20° towards the right), and the remaining noise dots are randomly oriented. Stimuli have been simplified (by reducing the number and size of dots) for illustration purposes.

Experiment 2 - Local and Global Motion

For the motion condition stimuli were presented for 750ms, and moved a distance of 3.2’ in each frame. For the local task all the dots moved upwards in the same direction at one of 9 angles either oriented on the vertical 0°, or tilted at one of the following angles ± 1.25°, 3.5°, 7.5°, 12.5°). Participants were required to indicate whether dots were moving left or right from the vertical mid-point. For the global task dots were randomly designated as signal or noise. Signal dots moved to the left or the right, noise dots moved in a random direction selected from 0-360°. Each dot maintained the same trajectory, but had a limited lifetime of 4 frames. The next trial only began after a response was made, preceded by a fixation cross for 500 ms, a further 500ms pause. The task here was to identify the direction (either left or right) of coherent motion provided by the signal dots.

Like the form task, test stimuli were presented before and after training, using MOCS and responses were made using the left or right arrows on the keyboard, also containing a randomised order of trials with 30 repetitions (270 trials). A cumulative Gaussian curve was fit to responses from the daily pre-assessment each day to determine a 65% and 85% correct discrimination threshold for use in the training blocks. The training task was conducted using the same QUEST methods outlined in the form task, using the motion stimuli.

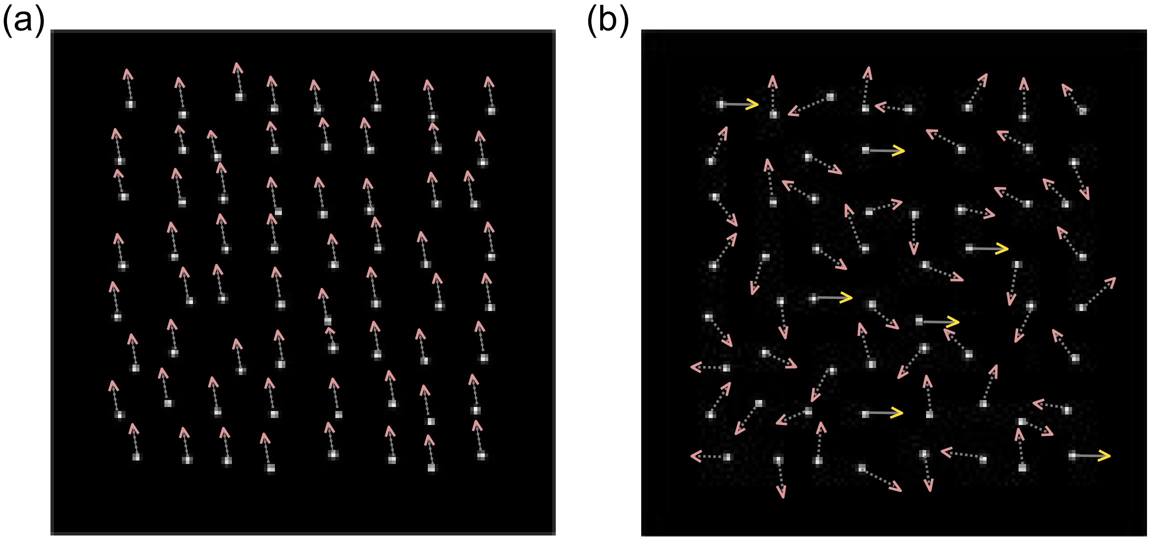


Figure. 5.4 Stimuli for the motion tasks; dots have been rendered larger and fewer in number for illustration purposes only. The arrows indicate the trajectory of motion. Left: Local motion stimuli where all dots are moving upwards and are oriented leftwards at 5° from the vertical. Right: Global motion stimuli where signal dots move rightwards and noise dots move in random directions.

Experiment 3 - Global Motion using MOCS training

For Experiment 3 pre- and post-assessments were identical to those used for experiment 2. However, to evaluate any training differences between methods of presentation (staircase or MOCS) these groups where trained using predefined randomly interleaved motion coherence levels. Training levels were identified by fitting a psychometric curve on the post-assessment results from the participants in the global motion condition from Experiment 2 to establish the mean number of dots required to achieve a percentage of correct responses ranging from 50% to 95% correct in increments of 5%. Coherence was defined at 8 levels (27, 28, 49, 78, 106, 135, 163, 210) which designated the number of signal dots present for that trial. Trials were presented randomly using MOCS for 10 repetitions at each signal level. Training consisted of 12 blocks with a total of 80 trials within each block, and 960 training trials daily.

5.3 Results

5.3.1 General Analysis information

Separate analyses were conducted for the training and the test data. For the local conditions (local form and local motion), scores were collapsed over the absolute stimulus values so that the nine ± levels left or right, became five absolute levels of degree of tilt. The results from the MOCS experiment are reported separately at the end of the results section.

Training - QUEST

For the training data, the threshold stimulus intensity producing a performance response level of 65% or 85% correct was determined. Training data were analysed using a general linear mixed model, conducted using Matlab (maximum likelihood model). Each model-fit was assessed for goodness of fit, using random slopes or intercepts only, or their combination. Models were compared to evaluate which model produced the lowest model fit statistics (Moscatelli et al., 2012) using the following parameters as comparison.

$$\begin{aligned} \text{response} &\sim \text{group} + \text{day} + \text{group} \times \text{day} + (1|\text{observer}) \\ \text{response} &\sim \text{group} + \text{day} + \text{group} \times \text{day} + (1 + \text{group}|\text{observer}) \\ \text{response} &\sim \text{group} + \text{day} \ + \text{group} \times \text{day} + (1 + \text{day}|\text{observer}) \end{aligned}$$

The model providing the best fit was:

$$\text{response} \sim \text{group} + \text{day} + \text{group} \times \text{day} + (1|\text{observer}) \tag{5.1}$$

Response (threshold) was modelled as a fixed effect of day and group (which was a categorical variable), and an interaction between these two predictors. The model also used random intercepts and slopes across participants.

Pre and Post-Assessment - MOCS

For all testing conditions, we evaluated the change in the proportion of correct responses across all stimulus intensities over all six sessions. Participants completed a pre- and post-training session on each of the three days, producing six sessions in total. Pre- and post-sessions were evaluated in sequence with numbers 1,3 and 5 the pre test, and sessions 2,4,6 the post tests for days 1,2 and 3 respectively. For each level of processing (local or global) the groups trained with and without feedback were independently analysed to evaluate if performance (the degree of tilt) improved across the six sessions (daily pre- and post-test). Again each model-fit was assessed for goodness of fit. The following combination of parameters were compared to evaluate which model produced the lowest model fit statistics (Moscatelli et al., 2012).

$$\begin{aligned}
\text{response} &\sim \text{session} + \text{tilt} + \text{session} \times \text{tilt} + (1|\text{observer}) \\
\text{response} &\sim \text{session} + \text{tilt} + \text{group} \times \text{tilt} + (1 + \text{session}|\text{observer}) \\
\text{response} &\sim \text{session} + \text{L}\ \ + \text{session} \times \text{tilt} + (1 + \text{tilt}|\text{observer})
\end{aligned}$$

Test data were then analysed using a general linear mixed model, conducted using Matlab (maximum likelihood model). Response (proportion correct) was modelled as a fixed effect of session and tilt and an interaction between these two predictors using with the following syntax:

$$\text{response} \sim \text{session} + \text{tilt} + \text{session} \times \text{tilt} + (1|\text{observers}) \tag{5.2}$$

Finally, the data sets were the combined to compare group differences. The proportion of correct responses was modelled as a fixed effect of group (feedback or no feedback), session (1-6) and tilt (stimulus intensity) as well as and an interaction term between all three fixed effects. The model also used random intercepts and slopes across participants with the following Matlab syntax:

$$\text{response} \sim \text{group} \times \text{session} + \text{group} \times \text{tilt} + \text{group} \times \text{session} \times \text{tilt} + (1|\text{observers}) \quad (5.3)$$

5.3.2 Experiment 1 - Local and Global Form

Training

The responses for local see Table 5.1 and global Table 5.2 groups' performance at 65% (difficult trials) and 85% (easy trials) were each analysed independently. Threshold was modelled as defined in Equation 5.1. Each group consisted of 14 observers who completed 12 blocks of training each day for three consecutive days. All participants attended every day. A negative coefficient (β) indicate a reduction in threshold and thus improvement, conversely a positive results indicates an increase in threshold and therefore worse performance. For the local form task for groups both with and without feedback thresholds improved significantly for both easy and difficult trials. There was a significant difference in both overall performance and the rate of improvement in threshold × day between groups.

The responses for the global form condition for 65% (difficult trials) and 85% (easy trials) were each analysed independently and are reported in Table 5.2. For the global form task for both groups (with or without feedback) thresholds improved significantly at both fixed levels.

Pre- and post-training tests

Assessment data were analysed independently for local and global form tasks and modelled using the previously defined Equation 5.2 for the within group analysis. An analysis that

Table 5.1 Local form training results across the three days for the measured thresholds at 65% and 85% accuracy, where the No-Feedback group are the reference group

Name	β	SE	tStat	DF	pValue	Lower CI	Upper CI
Local Form Threshold65 - Threshold							
No Feedback	1.7663	0.17512	10.086	1004	< 0.001	1.4226	2.1099
No Feedback Change	-0.18812	0.040432	-4.6527	1004	< 0.001	-0.26746	-0.10878
Feedback	1.2134	0.24766	4.8995	1004	$< 0.001^*$	0.72743	1.6994
Feedback Change	-0.31965	0.05718	-5.5903	1004	$< 0.001^*$	-0.43186	-0.20745
Local Threshold85 Form Threshold							
No Feedback	4.8428	0.45649	10.609	1004	< 0.001	3.9471	5.7386
No Feedback Change	-0.61666	0.069095	-8.9248	1004	$< 0.001^*$	-0.75225	-0.48107
Feedback	3.2753	0.64557	5.0736	1004	$< 0.001^*$	2.0085	4.5422
Feedback Change	-0.88295	0.097715	-9.0359	1004	$< 0.001^*$	-1.0747	-0.6912

Table notes :* indicates significant p value

Table 5.2 Global form training results across the three days for the measured thresholds at 65% and 85% accuracy, where the No-Feedback group are the reference group

Name	β	SE	tStat	DF	pValue	Lower CI	Upper CI
Global Form Threshold65 - Threshold							
No Feedback	58.806	5.094	11.544	1040	< 0.001	48.811	68.802
No Feedback Change	-8.0589	0.7481	-10.773	1040	$< 0.001^*$	-9.5269	-6.591
Feedback	5.6664	7.2161	0.78524	1040	0.432	-8.4934	19.826
Feedback Change	-3.6762	1.0767	-3.4143	1040	$< 0.001^*$	-5.7889	-1.5634
Global Threshold85 Coherence Threshold							
No Feedback	133.67	8.3143	16.078	1040	< 0.001	117.36	149.99
No Feedback Change	-13.568	0.87492	-15.508	1040	$< 0.001^*$	-15.285	-11.851
Feedback	15.809	11.768	1.3434	1040	0.179	-7.2829	38.902
Feedback Change	-9.364	1.2592	-7.4364	1040	$< 0.001^*$	-11.835	-6.8931

Table notes :* indicates significant p value included between groups performance was also modelled using the Matlab syntax defined in Equation 5.3. Local form results are reported in Table 5.3. Global form results are reported in Table 5.4. The test results measure the proportion of correct responses at each stimulus level, and a positive results indicates an improvement in performance; conversely a negative coefficient (β) indicates worsened performance.

For the *local form conditions* both groups showed a statistically significant improvement, both with and without feedback. The pattern of improvement differed between groups. The No-Feedback group showed no improvement to the intercept across the sessions, suggesting that there was no improvement at the lowest angles of tilt. In contrast the group who trained in the absence of feedback had a significant change to the intercept, suggesting an improvement in the accuracy of responses at the lowest stimulus intensities. Performance increased for both groups when stimulus intensity (angle of tilt) increased. However, the No-Feedback group showed no significant interaction between stimulus intensity and session of training. Performance increased uniformly across all stimulus levels. In contrast, the group trained with feedback did have a significant interaction between stimulus intensity and session of training, showing a significant improvement at higher levels of stimulus intensity, reflected by the steeper slope. The results are plotted in Figure 5.5 (a). For the comparison of both groups, the No-Feedback group was used as the comparison group. There was a significant main effect of Session, but the Group × Session interaction was not significant, and neither was the three way Group × Session × Tilt interaction. Both groups improved significantly over the six sessions, and this was not different whether feedback was present or absent. The Group × Tilt interaction was significant, as indicated by the negative coefficient (β). This indicates that collapsed across sessions the psychometric function for the stimulus intensity was steeper for the No-Feedback group and shallower for the feedback group. Overall the No-Feedback group produced more correct responses.

Table 5.3 Local Form across session comparison of proportion of correct responses across stimulus intensity (angle of tilt) between the Feedback groups. The No-Feedback group are the reference group

Day	β	SE	tStat	DF	pValue	Lower CI	Upper CI
No Feedback Group							
Session	-2.1005	0.17137	-12.257	416	< 0.001	-2.437	-1.764
Session (change)	0.12941	0.056709	2.282	416	0.023*	0.018	0.241
Tilt (degree)	0.7361	0.029823	24.683	416	$< 0.001^{\dagger}$	0.677	0.795
Session × Tilt	0.00846	0.018742	0.45126	416	0.652	-0.028	0.045
Feedback Group							
Session	-2.1413	0.20693	-10.348	416	< 0.001	-2.548	-1.735
Session (change)	-0.0252	0.063795	-0.39461	416	0.693	-0.151	0.100
Tilt (degree)	0.56353	0.029376	19.183	416	$< 0.001^{\dagger}$	0.506	0.621
Session × Tilt	0.05332	0.018577	2.8703	416	0.004*	0.017	0.090
Both Groups							
Group	-0.0348	0.2684	-0.1296	832	0.8969	-0.5616	0.4920
Session	0.1297	0.0567	2.2863	832	< 0.001*	0.0184	0.2411
Tilt	0.7368	0.0298	24.6950	832	$< 0.001^{\dagger}$	0.6782	0.7954
Group × Session	-0.1550	0.0854	-1.8155	832	0.0698	-0.3225	0.0126
Group × Tilt	-0.1738	0.0419	-4.1519	832	< 0.001*	-0.2560	-0.0916
Session × Tilt	0.0084	0.0187	0.4499	832	0.6529	-0.0284	0.0452
Grp × Session × Tilt	0.0448	0.0264	1.6974	832	0.0900	-0.0070	0.0966

Table notes :* indicates significant p value
and † performance increases significantly with increasing stimulus intensity (angle of tilt)

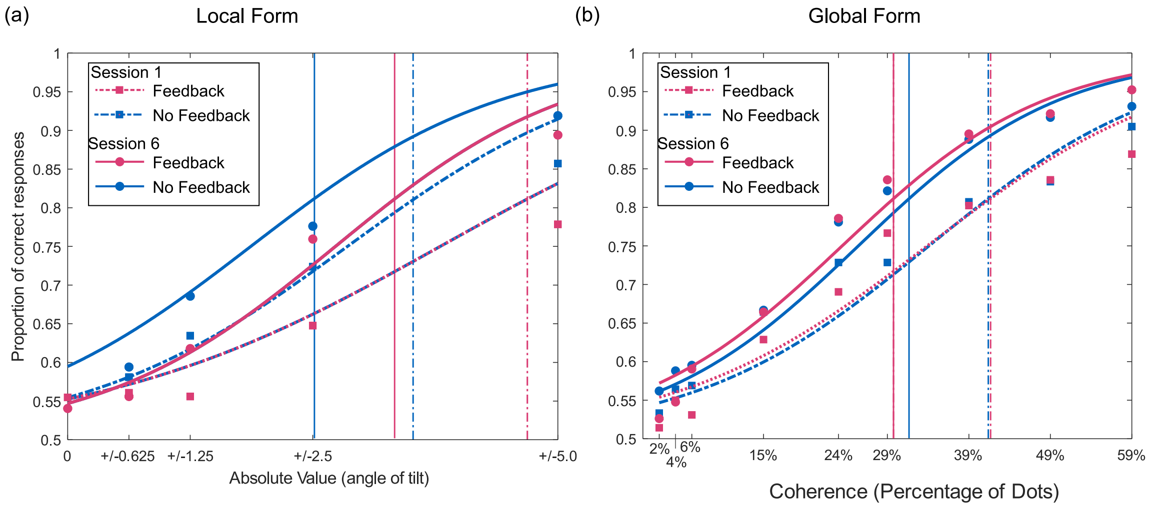


Figure. 5.5 Plots for (a) local and (b) global form where groups without feedback are plotted in pink and groups without feedback are plotted in blue. While the regression contained all sessions, performance has only been plotted for the first pre-test and for the final post-test (sessions 1 and 6 respectively). The local form plot (a) shows the local form The mean proportion of correct responses at each level of stimulus intensity (absolute value of angle of tilt). The global form plot (b) shows the mean proportion of correct responses as a function of the percentage of coherently oriented dots. The vertical lines mark the pre- and post (solid and dotted respectively) 75% thresholds, where the group with Feedback is in pink and No Feedback in blue.

For the *global form conditions*, again both groups showed a statically significant improvement across sessions, with and without feedback. The pattern of improvement was the same between groups. Both groups showed a significant improvement in intercept (Session), slope (Coherence) and a significant interaction between the two. There was no two-way interaction between groups across the sessions, nor was there a difference in the proportion correct obtained at stimulus intensities. The three-way interaction indicated no significant differences across groups. This shows that accuracy improved across all stimulus intensities, and there was no between groups difference. The results have been plotted in Figure 5.5 (b).

Table 5.4 Global Form: Comparison of proportion of correct responses across stimulus intensity (Coherence), by session within group, and between the Feedback groups, where the No-Feedback group are the reference group

Day	β	SE	tStat	DF	pValue	Lower CI	Upper CI
No Feedback Group							
Session (intercept)	-2.4081	0.26668	-9.0296	752	< 0.001	-2.9316	-1.8845
Session (change)	0.056844	0.01891	3.0053	752	0.003^{*}	0.0197	0.093975
Coherence	0.027448	0.00085	32.281	752	$< 0.001^{\dagger}$	0.025779	0.029117
Session × Coherence	0.000939	0.00022	4.3395	752	$< 0.001^{*}$	0.000514	0.00136
Feedback Group							
Session (intercept)	-2.2444	0.15972	-14.052	752	< 0.001	-2.5579	-1.9308
Session (change)	0.060897	0.018916	3.2194	752	0.0014^{*}	0.023763	0.098031
Coherence	0.02574	0.00084	30.781	752	$< 0.001^{\dagger}$	0.024099	0.027382
Session × Coherence	0.0012007	0.00022	5.5446	752	$< 0.001^{*}$	0.00078	0.0016
Both Groups							
Group	0.1521	0.3101	0.4906	1504	0.6238	-0.4562	0.7605
Session	0.0570	0.0189	3.0165	1504	0.0026^{*}	0.0199	0.0941
Coherence	0.0274	0.0008	32.2790	1504	$< 0.001^{\dagger}$	0.0257	0.0291
Group × Session	0.0040	0.0267	0.1496	1504	0.8811	-0.0485	0.0565
Group × Coherence	-0.0016	0.0012	-1.3601	1504	0.1740	-0.0040	0.0007
Session × Coherence	0.0009	0.0002	4.3131	1504	$< 0.001^{*}$	0.0005	0.0014
Grp × Sesh × Coh	0.0003	0.0003	0.8874	1504	0.3750	-0.0003	0.0009

Table notes :* indicates significant p value
and † performance increases significantly with increasing stimulus intensity (angle of tilt)

5.3.3 Experiment 2 - local and global motion

Training

Local and global groups' performance at 65% (difficult trial) and 85% (easy trial) were analysed separately. Threshold was modelled as reported in Equation 5.1 and the results are reported in Tables 4.4 (a & b). Each group consisted of 14 observers who all attended every session, however 1 observer dropped out of the local motion feedback condition, thus data were only analysed for the remaining 13 observers. Thus in the local analysis there were 972 data-points (individual threshold tallies, across 13 observers and 2 stimulus levels) and 4 fixed effects parameters (intercept, threshold, day and their interaction), from which the degrees of freedom were calculated.

A negative coefficient for the Change (β) indicates a reduction in threshold and thus improvement, conversely a positive coefficient for the Change (β) indicates in increase in threshold and therefore worse performance. Both groups (feedback and no-feedback) showed a reduction in threshold across the three days. Furthermore feedback, whether present or absent, had no effect on learning for any condition. For the local threshold at 85% correct there was a significant interaction between feedback and no feedback groups over the three days, with the No-Feedback group showing lower thresholds. (see Tables 4.4 (a))

Table 5.5 Local motion training results across the three days for the measured thresholds at 65% and 85% accuracy, where the No-Feedback group are the reference group

Name	β	SE	tStat	DF	pValue	Lower CI	Upper CI
Local Threshold65							
Day (intercept)	2.2965	0.18847	12.185	968	< 0.001	1.927	2.666
Day (change)	-0.19344	0.031715	-6.0992	968	$< 0.001^*$	-0.256	-0.131
Feedback (0 -1)	-0.27683	0.27161	-1.0192	968	0.308	-0.810	0.256
Day × Feedback	0.054724	0.045706	1.1973	968	0.231	-0.035	0.144
Local Threshold85							
Day (intercept)	5.667	0.39747	14.258	968	< 0.001	4.887	6.447
Day (change)	-0.54664	0.055555	-9.8396	968	$< 0.001^*$	-0.656	-0.438
Feedback (0 -1)	-0.9369	0.57282	-1.6356	968	0.102	-2.061	0.187
Day × Feedback	0.36775	0.080063	4.5933	968	$< 0.001^*$	0.211	0.525

Table notes :* indicates significant p value

All 14 observers attended every session for the global motion condition providing 1008 data-points (individual threshold tallies, across 14 observers and 2 stimulus levels) and 4 fixed effects parameters (intercept, threshold, day and their interaction), from which the degrees of freedom were calculated. Both groups (feedback and no-feedback) showed a reduction in threshold across the three days. Furthermore feedback, whether present or absent, had no effect on learning for any condition. The presence or absence of feedback did not impact learning at either level. (see Tables 4.4 (b))

Table 5.6 Global motion training results across the three days for the measured thresholds at 65% and 85% accuracy, where the No-Feedback group are the reference group

Name	β	SE	tStat	DF	pValue	Lower CI	Upper CI
Global Threshold65							
No Feedback	132.36	15.216	8.6988	1004	< 0.001	102.500	162.220
No Feedback Change	-22.403	1.9734	-11.352	1004	$< 0.001^*$	-26.275	-18.530
Feedback	-13.401	21.519	-0.62273	1004	0.534	-55.628	28.827
Feedback Change	2.3565	2.7909	0.84436	1004	0.399	-3.120	7.833
Global Threshold85							
No Feedback	158.81	15.923	9.9735	1004	< 0.001	127.560	190.060
No Feedback Change	-7.0676	1.6029	-4.4093	1004	$< 0.001^*$	-10.213	-3.922
Feedback	7.7245	22.519	0.34302	1004	0.732	-36.465	51.914
Feedback Change	-3.8988	2.2668	-1.72	1004	0.086	-8.347	0.549

Table notes :* indicates significant p value

Pre- and post-training tests

Data were assessed independently for local and global motion tasks, modelled using the previously reported Equation 5.2 for within group analysis, and then following this between groups performance was modelled using the Matlab syntax reported in Equation 5.3. Local motion results are reported in Table 5.7, and global motion tasks report in Table 5.8.

For the test results which measure the proportion of correct responses at each stimulus level, a positive coefficient for the slope (β) indicates an increase in performance, and conversely a negative coefficient for the slope (β) indicates worsened performance.

In the *local motion condition*, there was significant learning for the No-Feedback group over the six sessions, there was also a significant Session × Tilt interaction which indicates an increase in the number of correct responses as a function of stimulus levels between the

Table 5.7 Local motion Test: Comparison of proportion of correct responses across stimulus intensity (Degree of Tilt), by session within group, and between the Feedback groups, where the No-Feedback group are the reference group

Day	β	SE	tStat	DF	pValue	Lower CI	Upper CI
No Feedback Group							
Session 1	-1.536	0.24303	-6.3202	416	< 0.001	-2.014	-1.058
Session (change)	0.039437	0.014864	2.6531	416	0.008^*	0.01	0.069
Tilt	0.45737	0.012255	37.321	416	$< 0.001^\dagger$	0.433	0.481
Session × Tilt	0.0088279	0.0031829	2.7735	416	0.006^*	0.003	0.015
Feedback Group							
Session 1	-1.2984	0.19478	-6.666	386	< 0.001	-1.681	-0.915
Session (change)	-0.017499	0.015724	-1.1129	386	0.266	-0.048	0.013
Tilt	0.47951	0.01333	35.972	386	$< 0.001^\dagger$	0.453	0.506
Session × Tilt	0.00061	0.003366	0.16427	386	0.87	-0.006	0.007
Both Groups							
Group	0.2314	0.3140	0.7369	802	0.4614	-0.3850	0.8478
Session	0.0394	0.0149	2.6543	802	0.008^*	0.0103	0.0686
Tilt	0.4568	0.0122	37.2990	802	$< 0.001^\dagger$	0.4328	0.4808
Group × Session	-0.0570	0.0216	-2.6337	802	0.008^*	-0.0995	-0.0145
Group × Tilt	0.0233	0.0181	1.2873	802	0.1984	-0.0122	0.0589
Session × Tilt	0.0088	0.0032	2.7664	802	0.005^*	0.0026	0.0150
Grp × Sesh × Tilt	-0.0082	0.0046	-1.7788	802	0.0757	-0.0173	0.0009

Table notes :* indicates significant p value
and † performance increases significantly with increasing stimulus intensity (angle of tilt)

sessions. There was no improvement for the feedback group. The between groups analysis confirmed that the Group × Session interaction was significant, with the Feedback group performing significantly worse than the No-Feedback group (evident by the negative β). The three-way interaction between Group, session and tilt was not significant. The results are plotted in Figure 5.6 (a).

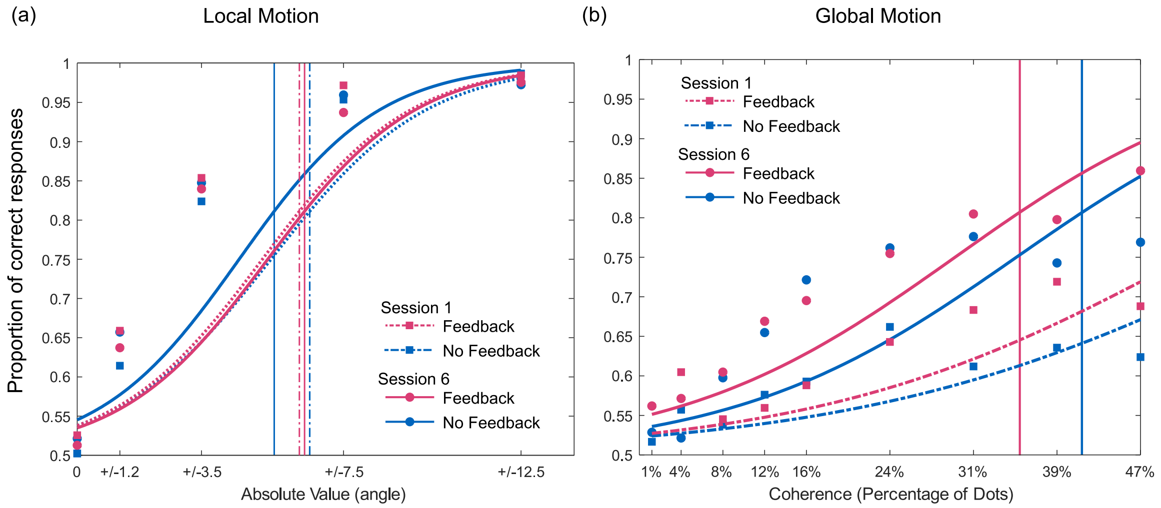


Figure. 5.6 Plots for (a) local and (b) global motion where groups without Feedback are plotted in pink and groups with No Feedback are plotted in blue. While the regression contained all sessions, performance has only been plotted for the first pre-test and for the final post-test (sessions 1 and 6 respectively). The local motion plot (a) shows average proportion of correct responses at each level of stimulus intensity (absolute value of angle of motion). The authors acknowledge that the fit is not ideal. The global motion plot (b) shows the mean proportion of correct responses as a function of the percentage of coherently moving dots. The vertical lines mark the post (dotted) 75% thresholds, where the group with Feedback is in pink and No Feedback in blue. The pre-training 75% thresholds for session 1 were beyond the scope of the plot.

Table 5.8 Global motion test: Comparison of proportion of correct responses across stimulus intensity (Degree of Tilt), by session within group, and between the Feedback groups, where the No-Feedback group are the reference group

	β	SE	tStat	DF	pValue	Lower CI	Upper CI
No Feedback Group							
Session (intercept)	-3.0481	0.505	-6.0414	752	< 0.001	-4.039	-2.056
Session (change)	0.08068	0.020191	3.9957	752	$< 0.001^*$	0.041	0.120
Coherence	0.01996	0.001083	18.431	752	$< 0.001^\dagger$	0.018	0.0221
Session × Coherence	0.001868	0.00027	6.9196	752	$< 0.001^*$	0.0013	0.0024
Feedback Group							
Session (intercept)	-2.9186	0.35468	-8.2289	752	< 0.001	-3.6149	-2.22
Session (change)	0.13217	0.0198	6.6751	752	$< 0.001^*$	0.093	0.171
Coherence	0.022251	0.0011	20.778	752	$< 0.001^\dagger$	0.0201	0.024
Session × Coherence	0.001531	0.0003	5.8315	752	$< 0.001^*$	0.001	0.002
Both Groups							
Group	0.1032	0.6063	0.1702	1504	0.8648	-1.0861	1.2925
Session	0.0808	0.0202	4.0014	1504	$< 0.001^*$	0.0412	0.1203
Coherence	0.0200	0.0011	18.4370	1504	$< 0.001^\dagger$	0.0178	0.0221
Group × Session	0.0514	0.0283	1.8173	1504	0.0694	-0.0041	0.1069
Group × Coherence	0.0023	0.0015	1.5257	1504	0.1273	-0.0007	0.0053
Session × Coherence	0.0019	0.0003	6.8974	1504	$< 0.001^*$	0.0013	0.0024
Grp × Sesh × Coh	-0.0003	0.0004	-0.8758	1504	0.3813	-0.0011	0.0004

Table notes :* indicates significant p value
and † performance increases significantly with increasing stimulus intensity (Coherence)

While neither groups in the *global motion* condition reached an asymptotic performance of 1, there was significant learning for both groups across the six sessions and both groups had more correct responses after training. Both groups also showed a significant interaction between session and coherence, which indicates that the proportion of correct responses over the sessions also improved as a function of coherence. While the Feedback group produced more correct responses at each level of stimulus intensity, this was not a significant difference and the three-way interaction between Group, Session and Coherence was not significant. Both groups improved similarly with or without feedback. The results are plotted in Figure 5.6 (b).

5.3.4 Experiment 3 - Global motion using MOCS

Each group (feedback or no feedback) consisted of 14 observers who all attended every session, with the exception of 1 observer who withdrew from the motion with feedback condition, thus data were only analysed for the remaining 13 observers. Global motion was the only condition conducted. As training methods were delivered using MOCS, the 85% and 65% thresholds were not tracked as in the QUEST conditions, for this reason analysis was undertaken using a similar format as that used to analysis the pre- and post-training assessments. First, assessment data was modelled independently for each group using the following Matlab syntax:

$$\text{proportion correct} \sim \text{day} + \text{coherence} + \text{day} \times \text{coherence} + (1|\text{observers}) \tag{5.4}$$

The proportion of correct responses was modelled as a fixed effect of *day* and *coherence* (stimulus intensity) and an interaction term between the two factors. To compare the performance between feedback and no feedback groups (see Equation 5.5, proportion correct was modelled as a fixed effect of *group* and *coherence* (stimulus intensity) with an interaction between the predictors. The models also used random intercepts and slopes, across participants.

$$\text{proportion correct} \sim \text{group} + \text{day} + \text{coherence} + \text{group} \times \text{day} \times \text{coherence} + (1|\text{observers}) \tag{5.5}$$

Training

The results for the groups trained using MOCS are reported in Table 5.9. Both groups showed significant improvement throughout the three days, however there was a significantly better improvement for the group trained with feedback,especially at the lower coherence levels (indicated by the significant main effect for the intercept day). The between groups comparison suggested that there was a significant difference between Groups across both days and Coherence, thus the Feedback group performed increasingly better over each level of coherence that the no feedback group, and also better than the No-Feedback group over the three days. The three-way interaction was not significantly different when taking all training factors into account.

Table 5.9 Global Motion Training results (MOCS). Comparison of proportion of correct responses across stimulus intensity (Coherence), by session within group, and between the Feedback groups, where the No-Feedback group are the reference group

	β	SE	tStat	DF	pValue	Lower CI	Upper CI
No Feedback Group							
Day1 (intercept)	-2.1275	0.46418	-4.5832	4028	< 0.001	-3.0375	-1.2174
Day (change)	-0.0524	0.028617	-1.8313	4028	0.067	-0.10851	0.0037
Coherence	0.01172	0.00051	23.052	4028	$< 0.001^{\dagger}$	0.0107	0.0127
Day × Coherence	0.0031	0.000241	12.784	4028	$< 0.001^{*}$	0.00261	0.0036
Feedback Group							
Day 1 (intercept)	-1.843	0.32864	-5.6079	3740	< 0.001	-2.4873	-1.1986
Day (change)	0.13109	0.027642	4.7424	3740	$< 0.001^{*}$	0.0769	0.1853
Coherence	0.01105	0.000496	22.277	3740	$< 0.001^{\dagger}$	0.0101	0.012
Day × Coherence	0.0029	0.0002	12.388	3740	$< 0.001^{*}$	0.002431	0.0033
Both Groups							
Group	0.2783	0.5749	0.4840	7768	0.6284	-0.8486	1.4051
Day	-0.0521	0.0286	-1.8216	7768	0.0685	-0.1082	0.0040
Coherence	0.0117	0.0005	23.0610	7768	$< 0.001^{\dagger}$	0.0107	0.0127
Group × Day	0.1832	0.0398	4.6040	7768	$< 0.001^{*}$	0.1052	0.2612
Group × Block	-0.0007	0.0007	-0.9351	7768	0.3498	-0.0021	0.0007
Day × Coherence	0.0031	0.0002	12.7700	7768	$< 0.001^{*}$	0.0026	0.0035
Grp × Day × Coh	-0.0002	0.0003	-0.5431	7768	0.5871	-0.0008	0.0005

Table notes :* indicates significant p value
and † performance increases significantly with increasing stimulus intensity (Coherence)

Pre- and post-training tests

Assessment data were modelled using the previously reported Equation 5.2 for within group analysis, and then following this between groups performance was modelled using the matlab syntax reported in Equation 5.3. Results are reported in Table 5.10.

Within groups analysis found that the No-Feedback group did not improve significantly at the lower levels of coherence (no significant change for Day (Table 5.10). However, there was a significant session × coherence interaction, indicating better performance at increasing stimulus intensity, which also increased as a function of session. For the Feedback group improvement was robust across all stimulus intensities and this increased as a function of session.

The between groups analysis found no difference between group and coherence, suggesting that both groups improved as stimulus intensity increased. There was a significant group by session interaction, which suggests that collapsed across coherence the Feedback group were better than the No-Feedback group at detecting motion coherence at the lowest stimulus values after training. Finally there was a significant three-way interaction between Group, Session and Coherence; this suggests that there was a significant difference in the slope of the function (proportion correct as a function of stimulus intensity), across the sessions, between the groups - and while performance was better from the MOCS trained group, the slope of the function was steeper (more learning at higher levels of coherence) for the No-Feedback group across the sessions, compared to the feedback group.

While there was a significant improvement for both groups, the Feedback group in the MOCS condition showed more improvement across sessions at lower coherence while the No-Feedback group showed more improvement at higher coherence. The results are plotted in Figure 5.7(b) with the results from the QUEST test repeated in Figure 5.7 (a)

Table 5.10 Global Motion test (MOCS). Comparison of proportion of correct responses across stimulus intensity (Coherence), by session within group, and between the Feedback groups, where the No-Feedback group are the reference group

	β	SE	tStat	DF	pValue	Lower CI	Upper CI
No Feedback Group							
Session (intercept)	-3.7256	0.42113	-8.8467	752	< 0.001	-4.5523	-2.8988
Session (change)	-0.022075	0.03368	-0.6554	752	0.5123	-0.088	0.044
Coherence	0.023199	0.00152	15.255	752	$< 0.001^{\dagger}$	0.020	0.0262
Session × Coherence	0.0033804	0.00038	8.8504	752	$< 0.001^{*}$	0.0026	0.0041
Feedback Group							
Session (intercept)	-3.7031	0.29833	-12.412	698	< 0.001	-4.2888	-3.1173
Session (change)	0.12014	0.031979	3.757	698	$< 0.001^{*}$	0.0574	0.18293
Coherence	0.025338	0.00153	16.547	698	$< 0.001^{\dagger}$	0.022	0.0283
Session × Coherence	0.0021644	0.00036	5.9875	698	$< 0.001^{*}$	0.0015	0.0029
Both Groups							
Group	-0.0042	0.5175	-0.0081	1450	0.9935	-1.0194	1.0109
Session	-0.0218	0.0337	-0.6478	1450	0.5172	-0.0878	0.0442
Coherence	0.0232	0.0015	15.2540	1450	$< 0.001^{\dagger}$	0.0202	0.0262
Group × Session	0.1420	0.0465	3.0563	1450	0.0023^{*}	0.0509	0.2331
Group × Coherence	0.0022	0.0022	1.0199	1450	0.3079	-0.0020	0.0064
Session × Coherence	0.0034	0.0004	8.8402	1450	$< 0.001^{*}$	0.0026	0.0041
Grp × Sesh × Coh	-0.0012	0.0005	-2.2810	1450	0.0227^{*}	-0.0022	-0.0002

Table notes :* indicates significant p value
and † performance increases significantly with increasing stimulus intensity (Coherence)

MOCS vs QUEST

Test data from the respective conditions (MOCS or QUEST) were grouped by feedback, and analysed independently for comparison using the following Matlab syntax:

$$\text{proportion correct} \sim \text{coherence} \times \text{day} \times \text{condition} + (1 + \text{coherence}|\text{observers}) \qquad (5.6)$$

The proportion of correct responses was modelled as a fixed effect of *day* and *coherence* (stimulus intensity) and *condition* (MOCS or QUEST) and interactions term between the three factors. Random intercepts and slopes were used for participants and coherence and the data.

Table 5.11 MOCS vs QUEST - test comparison

	β	SE	tStat	DF	pValue	Lower CI	Upper CI
No Feedback (where conditions are Quest = 1, MOCS = 2)							
Session	-1.3007	0.5116	-2.5424	1504	0.0111	-2.3043	-0.2972
Session (change)	0.1455	0.0525	2.7732	1504	0.0056*	0.0426	0.2484
Coherence	0.0044	0.0082	0.5310	1504	0.5955	-0.0117	0.0204
Condition	-0.9508	0.3358	-2.8312	1504	0.0047*	-1.6096	-0.2921
Session × Coherence	0.0014	0.0007	2.0221	1504	0.0433*	< 0.001	0.0027
Session × Condition	-0.0899	0.0396	-2.2688	1504	0.0234*	-0.1677	-0.0122
Coherence × Condition	0.0068	0.0052	1.2958	1504	0.1952	-0.0035	0.0171
Sesh × Coh × Cond	0.0011	0.0005	2.2337	1504	0.0256*	0.0001	0.0020
Feedback (where conditions are Quest = 1, MOCS = 2)							
Session	-2.1165	0.7203	-2.9386	1450	0.0033	-3.5294	-0.7037
Session (change)	0.1442	0.0509	2.8329	1450	0.0047*	0.0443	0.2440
Coherence	0.0191	0.0026	7.2507	1450	$< 0.001^{\dagger}$	0.0139	0.0243
Condition	-0.7971	0.4684	-1.7017	1450	0.0890	-1.7160	0.1218
Session × Coherence	0.0009	0.0006	1.3670	1450	0.1719	-0.0004	0.0021
Session × Condition	-0.0120	0.0376	-0.3193	1450	0.7495	-0.0858	0.0618
Coherence × Condition	0.0031	0.0019	1.6805	1450	0.0931	-0.0005	0.0068
Sesh × Coh × Cond	0.0007	0.0004	1.4583	1450	0.1450	-0.0002	0.0015

Table notes :* indicates significant p value
and † performance increases significantly with increasing stimulus intensity (Coherence)

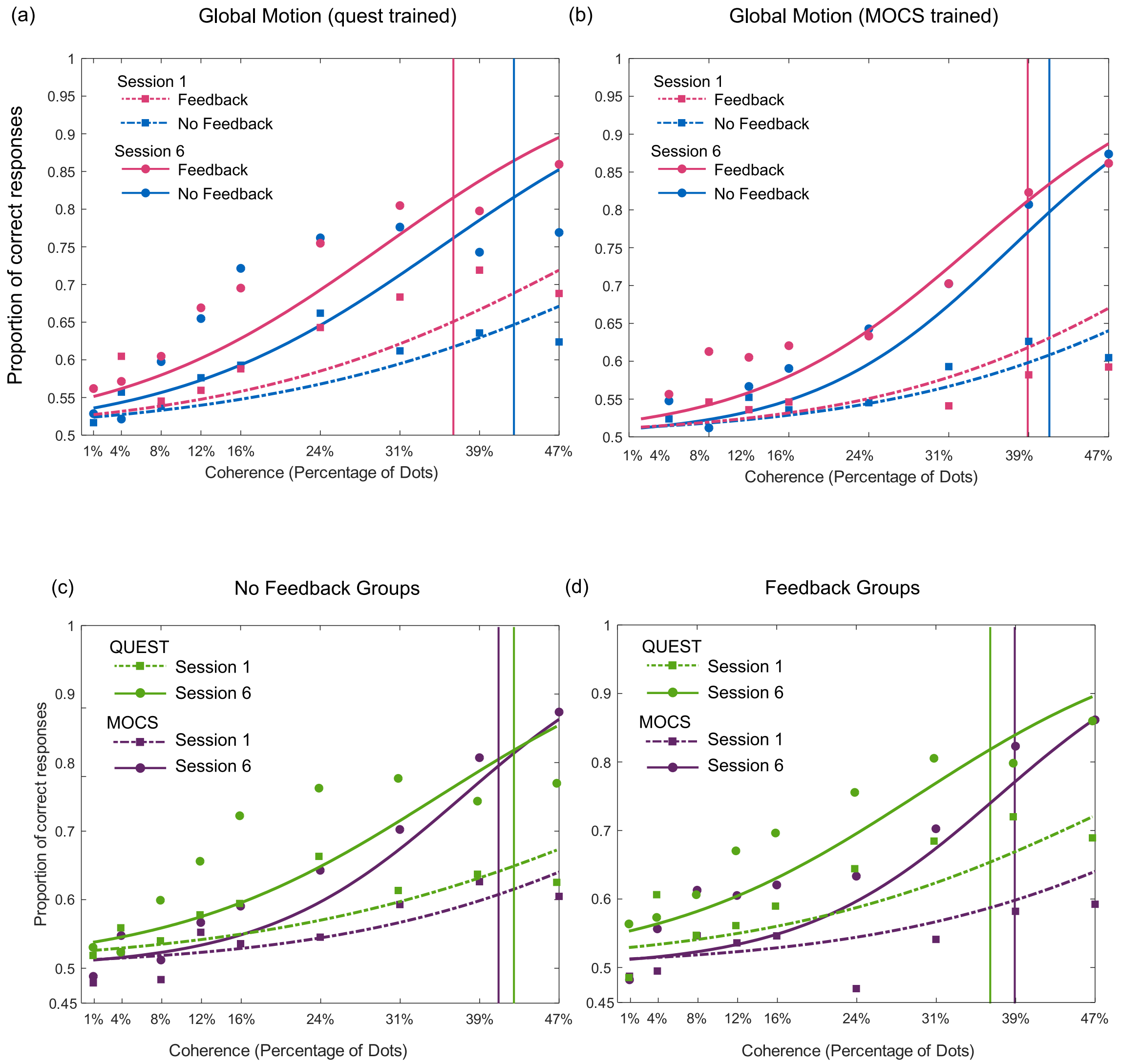


Figure. 5.7 Global Motion MOCS vs QUEST

For the *groups trained without feedback* there was a main effect of Session, (in general responses at lower coherences improved with training), and the Session by Condition interaction was significant. The QUEST trained group showed a significant improvement in the intercept, while the MOCS trained group did not. There was *no* main effect of coherence, which means that without feedback performance *did not* improve to a significantly higher degree with increasing intensity, and this was not different for the different training methods. There was a significant Session × Coherence interaction which suggests that the slope of the function was steeper for the MOCS trained group (the proportion of correct responses at increasing intensities. This is shown in Figure 5.7 (c) where at higher stimulus levels the MOCS group performed better than the QUEST group. Finally there was a significant three-way interaction between Session, Coherence and Condition indicating that across sessions, the change in the steepness of the slope of the psychometric function (proportion of correct responses across stimulus intensity increased more for the MOCS trained group than for the QUEST trained group.

For the *groups trained with feedback* There was a main effect of Session (in general responses at lower coherences improved with training), and no interaction between Session and Condition, indicating both groups improved at lower stimulus intensities (see Figure 5.7 (d)). Unlike the no feedback conditions, there was a significant main effect of Coherence whereby performance increased with higher stimulus intensities. There were no other significant differences between MOCS and QUEST trained groups with feedback.

5.4 Discussion

The primary aim of this study was to resolve some of the ambiguity in the findings of perceptual learning research that found (Liu et al., 2012a) or did not find (Asher et al., 2018b; Seitz et al., 2006) learning in a specifically designed training paradigm that interleaved easy and difficult trials to obtain learning without feedback. We considered a number of factors that may account for the differences in findings, namely (1) the level of processing (local or global) (2) the type of task (motion or form) (3) the absence or presence of external trial-by-trial feedback and (4) the type of training method (staircase or MOCS). We predicted that because of the stimulus driven nature of local tasks, we should secure robust improvement with and without feedback, evidenced during testing and training phases. At the global level, if there is a difference in how learning operates at higher level of segmentation and integration processing, compared to local processes, this would be evidenced in the no-feedback conditions of both (or either) motion and form tasks. If performance is contingent on the teaching signal provided by the external feedback (as reported for Experiment 2 in Chapter 4 (Asher et al., 2018b)), then this may be reflected by a reduction in learning during the testing conditions which contain no feedback. Finally, if presenting stimuli using a staircase provides bootstrapping to learning, while MOCS does not, we would predict that when using MOCS with no feedback learning would be reduced or absent.

Our results suggested robust learning across all conditions both with and without feedback, irrelevant of local or global motion or form tasks, or whether presented with a staircase or MOCS. However, while learning was a clear finding in general, it was not uniform across all conditions.

Experiment 1 - Local and Global Form

For *local form,* both groups (feedback and no feedback) showed a statistically significant improvement across the six testing sessions. The group trained with feedback showed a significant improvement at larger angles of tilt but no change at the lower angles of tilt. In contrast, the group trained with no feedback improved uniformly across all stimulus intensities, including at the lower stimulus intensities. Overall the no feedback group's performance was better (a higher proportion of correct responses) than those in the feedback group. However, improvement across groups, sessions and stimulus intensities was the same. For *global form*, again both groups showed a statistically significant improvement, with and without feedback improvement across all stimulus intensities was uniform, and there were no between groups differences.

Experiment 2 - Local and Global Motion

For *local motion,* While training (as reported in Table 4.4) was successful at decreasing the thresholds (65% and 85% correct) for both groups, only the No-Feedback group showed any improvement across the sessions. For *global motion,* while performance never reached perfect, learning was robust and consistent between the two groups.

Experiment 3 - Global Motion (MOCS Training)

The rationale for focusing only on the Global Motion condition with MOCS training was to resolve some of the unresolved questions from Chapter 4, where learning occurred for global motion, but did not persist at the testing phase for some conditions. The results suggest a complicated relationship between the factors.

Comparing the results from the feedback and no feedback groups :

- MOCS, no feedback, showed no improvement at lower coherence, but did at higher coherence.
- MOCS with feedback showed robust improvement at lower and higher coherence.
- At higher coherence the No-Feedback group performed better than the feedback group.
- At lower coherence the Feedback group performed better than the no feedback group.
- Across the sessions the proportion correct (across stimulus intensity) was higher for the feedback group.

Comparing the results from the MOCS and quest trained sub-groups:

No Feedback conditions

- Generally, there was no significant increase in performance across stimulus intensities (which is unusual), since asymptotic performance was never reached, this suggests the stimulus intensities did not cover a comprehensive enough range.
- In general QUEST trained group obtained a higher proportion of correct responses performing better than the MOCS trained group, especially at low stimulus intensities.
- However, a shift in the steepness of the slope from the No-Feedback group indicates more correct responses at higher intensities, and reaching a similar asymptote as the QUEST trained group.

Feedback conditions:

There were no significant differences in improvement for the MOCS or QUEST trained groups when feedback was present.

5.4.1 Predictions and Findings

Predictions based on previous findings and theoretical considerations were set out in the introduction. Here, we outline how our results compare with these predictions, and how these may be interpreted.

1. Local tasks should improve with or without feedback

Our first prediction that, given the nature of stimulus driven improvements at the local level, local tasks should improve with or without feedback. Our results showed that for both motion and form conditions i) learning took place and ii) overall there was no difference in performance and improvement with or without feedback. However, performance for the no feedback group *in the local form condition* significantly improved at smaller degrees of tilt compared to the feedback group. The change in slope of the psychometric function was steeper for the group with feedback and there was a large shift leftwards in the psychometric function, indicative of a decrease in threshold after training. There was no change in the steepness of the slope for the group trained with No Feedback, however the results also show a uniform leftward shift indicating improvement at all stimulus intensities. While there was a decrease in threshold for the No Feedback group, they were also lower than those of the feedback group, at the start of the experiment. *The local motion conditions* suggest that improvement was restricted to the No-Feedback group who provided an increased proportion of correct responses across all stimulus intensities and a rightward shift of the psychometric function (indicative of a lower threshold). The results indicate that as a function of Session

(1-6) the group with feedback had a significantly lower performance that the No Feedback at lower stimulus intensities.

Consistent in both local conditions, the groups with No Feedback showed a significant improvement in the intercept (performance at lower stimulus intensities) where there is more uncertainty in the stimulus. Neither group with Feedback had a significant change in intercept. Furthermore, the between groups analysis indicated that the change of intercept (Session) was significantly lower for Feedback groups when compared to No Feedback groups. This may suggest that for local tasks where ambiguity stems from reduction in stimulus intensity rather than ambiguity from added noise, the decision is more reliant on internally generated feedback. This may also suggest that at lower stimulus intensity the presence of trial-by-trial feedback may negatively interact with an observer's internal confidence, resulting in the undermining of internal confidence.

Generally, performance on the local motion task was better than performance on the form task. This is not surprising given the evolutionary importance of motion detection for behavioural action (Gibson, 2014). Furthermore, there would have been a high correlation in neural responses from local and global motion receptive fields, consistent with feedback from the re-entrant connections from V5 to V1 (Rockland & Knutson, 2000).

2. Differences between learning mechanisms at global levels of processing would be shown when comparing performance between Feedback conditions

If global levels are more reliant on top-down information in sensory processing and learning is contingent on the presence of feedback, this should be reflected in two places. Firstly, there might be a reduction in performance when feedback is removed (as found by Asher et al. (2018b)). Secondly, there may be steeper slopes for those trained with external feedback, which would suggest that global levels may be more reliant on external reinforcement for learning. In general performance for the global form tasks was better than for the global motion tasks. At the start of the experiment results suggest that performance on global motion tasks did not improve across stimulus intensity, suggesting that the task was very difficult and uncertainty was high. However, this improved by the end of the first day and over the course of the training.

However, importantly for both global conditions, learning was robust and consistent with or without trial-by-trial feedback. There were significant improvements in intercept and slope, for all global form and motion groups with no between groups differences for any measure. This suggests that internal and/or external feedback illicit similar improvements in performance in training, and learning was persistent even once feedback is removed.

3. Differences between learning mechanisms between global levels of motion and form conditions

Due to the different nature of each of the individual tasks no dedicated analysis was conducted between form and motion conditions. However, patterns of learning were robust, with and without feedback when easy trials are interleaved improved motion and form at global levels. This suggests that while processed in different areas of the visual cortex there is no difference in learning mechanisms that process form and motion. Supporting those theories that predict a higher level decision mechanism (Dosher & Lu, 2017).

4. Differences in learning as a result of the method of presentation

Our final prediction was that, if using an adaptive staircase acted as a ‘bootstrap’ to learning (Seitz et al., 2006), then in the conditions with No Feedback, QUEST trained groups would show higher levels of performance than those groups trained with MOCS. While the groups trained with Feedback showed no significant differences in their improvement, there were differences in the No Feedback trained groups.

In the Global Motion conditions where no feedback was presented, comparing post-training results between QUEST and MOCS trained groups, shows that QUEST trained observers improved at lower stimulus intensity, where the stimulus was ambiguous due to external noise. In contrast, the MOCS trained group showed no improvement at lower stimulus intensities (no change to the intercept), however as stimulus intensity increased, so did the proportion of correct responses, and at the highest intensity performance matched the performance of the QUEST trained group. For the groups trained with Feedback there was a significant improvement responses at lower stimulus intensities (change to the intercept) for both MOCS and QUEST trained groups. The change to the intercept was slightly lower for MOCS trained groups, however the Session by Condition interaction in Table 5.11 indicated this was not significantly so.

The change in the steepness of the slope of the psychometric function for the MOCS trained groups was significantly higher than the change in slope for the QUEST trained groups. Furthermore, this pattern was present across both the Feedback conditions, although it was not statistically significant for the groups trained with feedback. This shows that at higher stimulus intensities, MOCS trained groups performed better, resulting in a significant three-way interaction and better overall threshold for the MOCS trained group. The methods and levels for the testing phase was identical, only the training differed. An explanation for this may be that, as a result of the pre-defined training levels, the MOCS trained groups were exposed to a consistent range of difficulties in training. For the QUEST trained groups, training levels (coherence required in order to obtain a performance accuracy of 65% and 85%) were calculated daily based on their performance on the pre-test. Therefore their exposure to coherence intensities, required to maintain the defined performance level, reduced each day. In contrast, even by the final day the MOCS trained groups were still being trained on stimuli with extremely high coherence levels (up to 210 of 255 dots or 82% coherence). The QUEST group still exhibited significant effects of learning, however, this may suggest that reducing stimulus intensity as a function of performance may be less effective at generating robust learning, and that including a wider range of levels (whether using MOCS or staircasing) would likely be a more successful learning paradigm. Furthermore, at lower stimulus intensities QUEST methods elicited better performance (when there was No Feedback) suggestive of providing a ‘bootstrap’ to the decision reweighting that was not present for the MOCS group at the lowest intensities.

To improve using MOCS at low or sub-threshold coherences, methods that included either a staircase or trial-by-trial feedback was a requirement to obtain learning in this global motion task. This suggests that when there is a high level of external noise in the stimulus, the internal signal provided by the presence of easy exemplars, at set thresholds facilitated learning when trained with QUEST, but not when trained with randomly interleaved intensities using MOCS.

Chapter 6

First- and second-order contributions to depth perception in anti-correlated random dot stereograms

6.1 Introduction

Binocular depth perception depends on our ability to determine the difference in position of corresponding points between the two eyes' images. As outlined in the introduction, this occurs through a process of similarity matching. This is typically construed as a process of cross-correlation (Allenmark & Read, 2011; Cormack et al., 1991; Fleet et al., 1996; Julesz, 1971), although other metrics have also been proposed (Doi & Fujita, 2014; Doi et al., 2011).

These similarity calculations can be related to the responses of binocular neurons in the visual cortex (Allenmark & Read, 2011; Banks et al., 2004; Fleet et al., 1996; Nienborg et al., 2004). Neurons in V1 have a localised receptive field, consisting of both excitatory and inhibitory regions, and tend also to be tuned to orientation and spatial frequency (Cumming & Parker, 1997; Cumming et al., 1998; Fleet et al., 1996; Ohzawa et al., 1990, 1997). Binocular neurons have a receptive field in both eyes, and their responses are thus sensitive to binocular disparities. Figure 6.1 (a) shows an idealised binocular energy neuron, (Fleet et al., 1996; Ohzawa et al., 1990, 1997) consisting of a quadrature pair of receptive fields for each eye. The responses of the first-stage filters are the square of the sum across the two eyes' receptive field (Figure 6.1(b)), which in turn are summed to compute the binocular energy response.

These components of the energy model have been used to characterise the responses of binocular simple and complex cells, respectively. However, this simple hierarchy is best viewed as an idealised approximation of how these computations are performed (Martinez & Alonso, 2003). The response of a binocular energy neuron depends on the disparity in the images, forming a characteristic disparity tuning function (Figure 6.1(c)). In this example, the receptive fields of the filters are shifted horizontally in the right eye compared to those in the left eye, meaning that this neuron responds most strongly to stimuli with the same disparity. This peak response at this optimum disparity is accompanied by disparities at which the response is reduced compared to baseline.

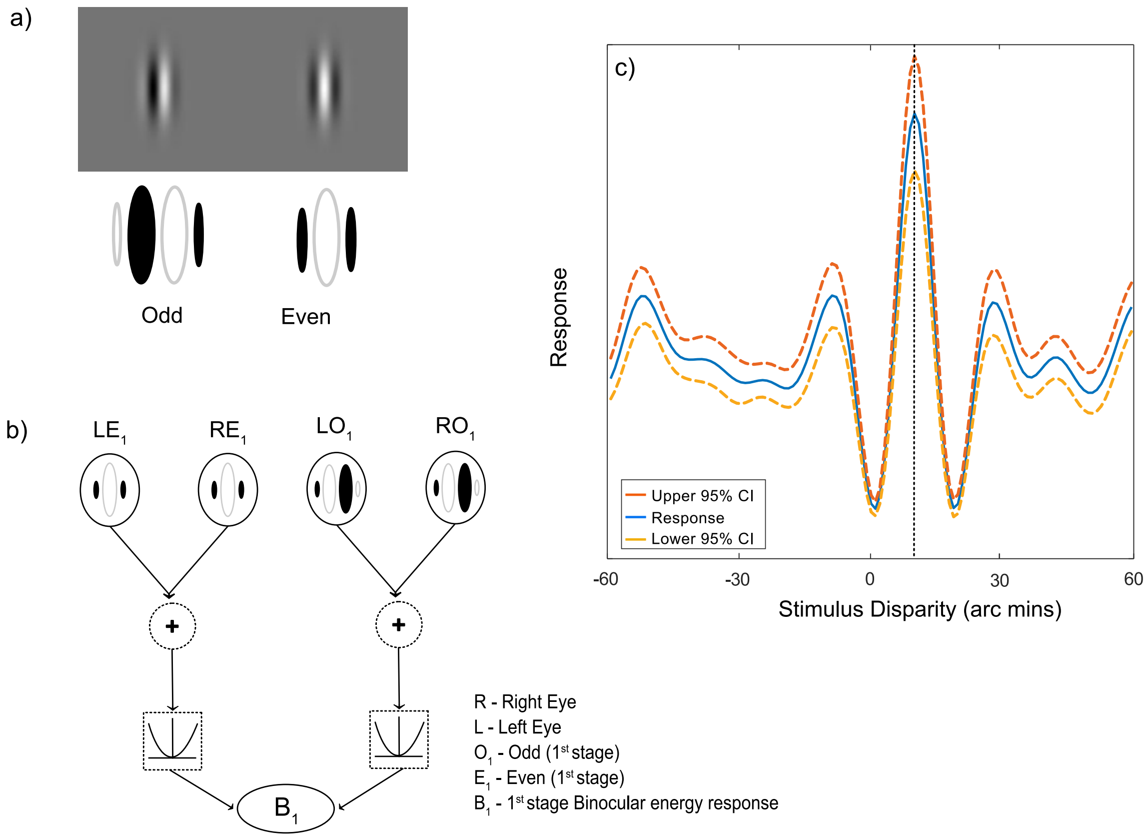


Figure. 6.1 (a) Odd and even symmetric monocular receptive fields (b) These receptive fields are combined and squared to compute the rectified binocular energy response (c) The responses of binocular energy neurons show a characteristic tuning function. Here the response, calculated as the mean over 1000 Gaussian white noise samples, shows a peak at the disparity (10 arc-min) corresponding to the difference in position of the monocular receptive fields. Dotted lines show upper and lower 95% confidence intervals.

The binocular energy response is related to the point-wise correlation between the filtered left and right images (Fleet et al., 1996). By summing across frequency, orientation and position (Fleet et al., 1996; Hibbard et al., 2016; Tanaka & Ohzawa, 2006), the correlation within a spatial neighbourhood can be calculated (Allenmark & Read, 2011; Fleet et al., 1996). Complex cells in V1, as characterised by the binocular energy model, are thus well suited to support the calculation of the cross-correlation and other matching calculations thought to contribute to the solution of the correspondence problem (Ohzawa et al., 1990).

The dependence of the energy response on binocular cross-correlation means that manipulation of this correlation has provided a useful way of understanding how the responses of populations of binocular neurons contribute to the perception of depth (Cormack et al., 1991). This is often investigated psychophysically using random-dot stereograms (RDS) which, by projecting an image that has been shifted slightly between the left and right eye, creates the perception of depth. Correlated random-dot stereograms (CRDS) present dots of the same luminance to each eye ((Figure 6.2 (a)). However, an interesting application of RDS is the use of anti-correlated random-dot stereogram stimuli (ACRDS), where one eye's view is replaced with its photographic negative (Figure 6.2(b)) (Anstis & Rogers, 1975; Aoki et al., 2017; Cogan et al., 1993; Cumming et al., 1998; Doi et al., 2011; Hibbard et al., 2014; Kaufman & Pitblado, 1965; Read & Eagle, 2000; Rogers & Anstis, 1975; Tanabe et al., 2008). This means that the high positive correlations expected at the correct disparity become negative, and the disparity tuning function is inverted. Neurons in V1 show this inversion effect, but also a reduction in magnitude of their response that is not predicted by the energy model (Cumming & Parker, 1997; Ohzawa et al., 1990, 1997). This reduction in response has been modelled using the introduction of a threshold non-linearity, (Lippert & Wagner, 2001; Nieder & Wagner, 2001; Read et al., 2002) or a squaring of the energy response (Henriksen et al., 2016). These expansive nonlinearities, by enhancing the difference between the amplitudes of the positive and negative peaks in the disparity tuning function, can be used to implement the cross-matching mechanism proposed by Doi and Fujita (Doi & Fujita, 2014; Doi et al., 2011).

In higher visual areas in the ventral stream, the responses of neurons tend not to be modulated by the disparity in ACRDS (Janssen et al., 2003; Tanabe et al., 2004). In contrast, neurons in dorsal stream areas show disparity tuning similar to that found in V1, but reduced in magnitude (Krug et al., 2004).

In some psychophysical studies, the direction of depth perceived in ACRDS has been found to be reversed in comparison with equivalent CRDS (Aoki et al., 2017; Doi et al., 2011; Read & Eagle, 2000; Tanabe et al., 2008). Thus, we use the term ‘reversed-depth’ to describe depth that is perceived in the opposite direction to the correct direction, and the term ‘forward-depth’ when it is perceived in the correct direction. One possible explanation of this percept is that it reflects the peaks in the inverted disparity tuning function. The signal for reversed- or forward-depth would depend on the relationship between the spatial frequency and disparity tuning of the neuron, and the stimulus disparity (Figure 6.3).

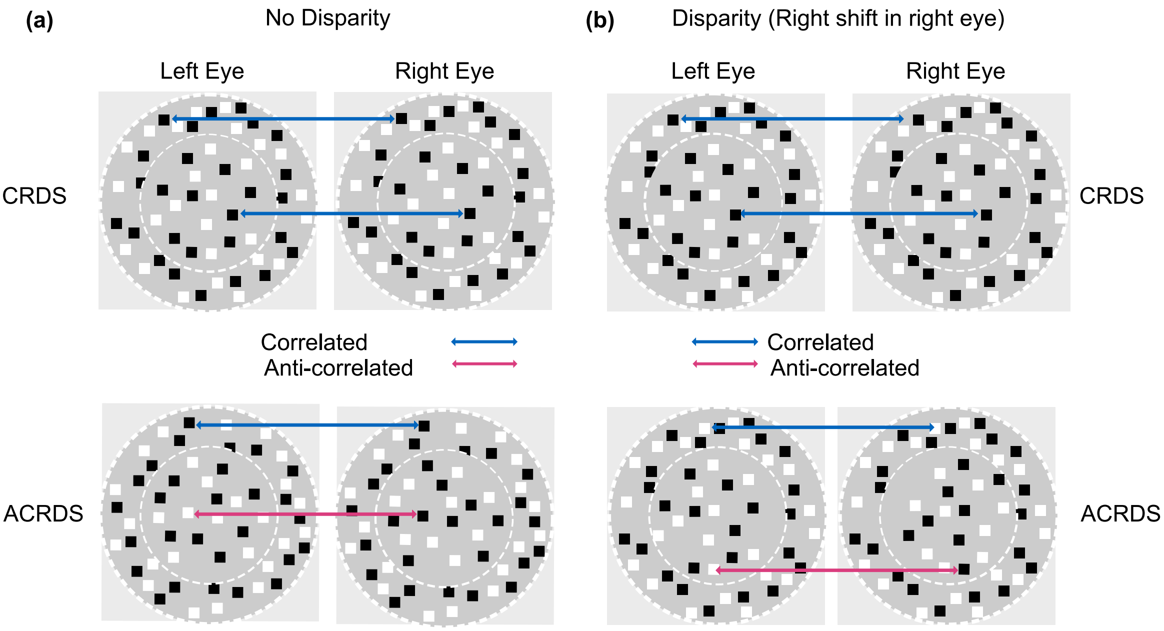


Figure. 6.2 (a) Correlated and Anti-correlated RDSs, left image is presented to the left eye and the right to the right eye. Both CRDS and ACRDS have a correlated reference (surround annulus). Correlated RDS (top) have correlated dots in the target (centre circle). In contrast, the central targets in anti-correlated RDS (bottom) have reversed luminance for the corresponding eyes. (b) Dots in the target (centre circle) include a rightwards shift for the right eye's image. The surround for all cases remains correlated with zero disparity.

Alternatively, it has been suggested that the estimation of depth might reflect opponent processing, in which the difference between the responses of neurons tuned to equal but opposite disparities are calculated (Doi & Fujita, 2014; Henriksen et al., 2016). In this case, the negative correlations that exist in ACRDS would directly contribute to the reduction in the response to the correct disparity, thereby biasing towards the perception of depth towards the opposite direction.

Other studies have found no evidence for the perception of depth in ACRDS (Cogan et al., 1993; Hibbard et al., 2014; Read & Eagle, 2000). This result is consistent with the fact that there is no coherent disparity, across different scales of analysis, in ACRDS, and also with the lack of disparity-selective responses in higher cortical areas.

The effect of decorrelation has been further assessed by Doi et al. (2011) who created stimuli containing an equal mixture of correlated and anti-correlated dot pairs. This results in an overall correlation of zero, so that a correlation-based mechanism predicts that depth would not be seen in these half-matched stimuli. In fact, depth is perceived in the correct direction. This is consistent with the responses of the cross-matching mechanism, which responds to the correct-matched dot pairs, but not to the anti-correlated pairs Henriksen et al. (2016) showed that a similar prediction can be made for the squared energy response, which enhances the difference between the response to the paired and unpaired dots.

Predicting the perception of depth in ACRDS is complicated by the fact that there is no coherent disparity signal across different spatial scales. In CRDS, pooling of information across frequency, scale and position allows true peaks to emerge from amongst the many large responses that will occur at incorrect disparity values (Fleet et al., 1996; Hibbard et al., 2016; Tanaka & Ohzawa, 2006). This process does not produce a clear estimate for ACRDS, since large peaks are predicted to occur at different locations at different spatial scales (Hibbard et al., 2014) (Figure 6.3).

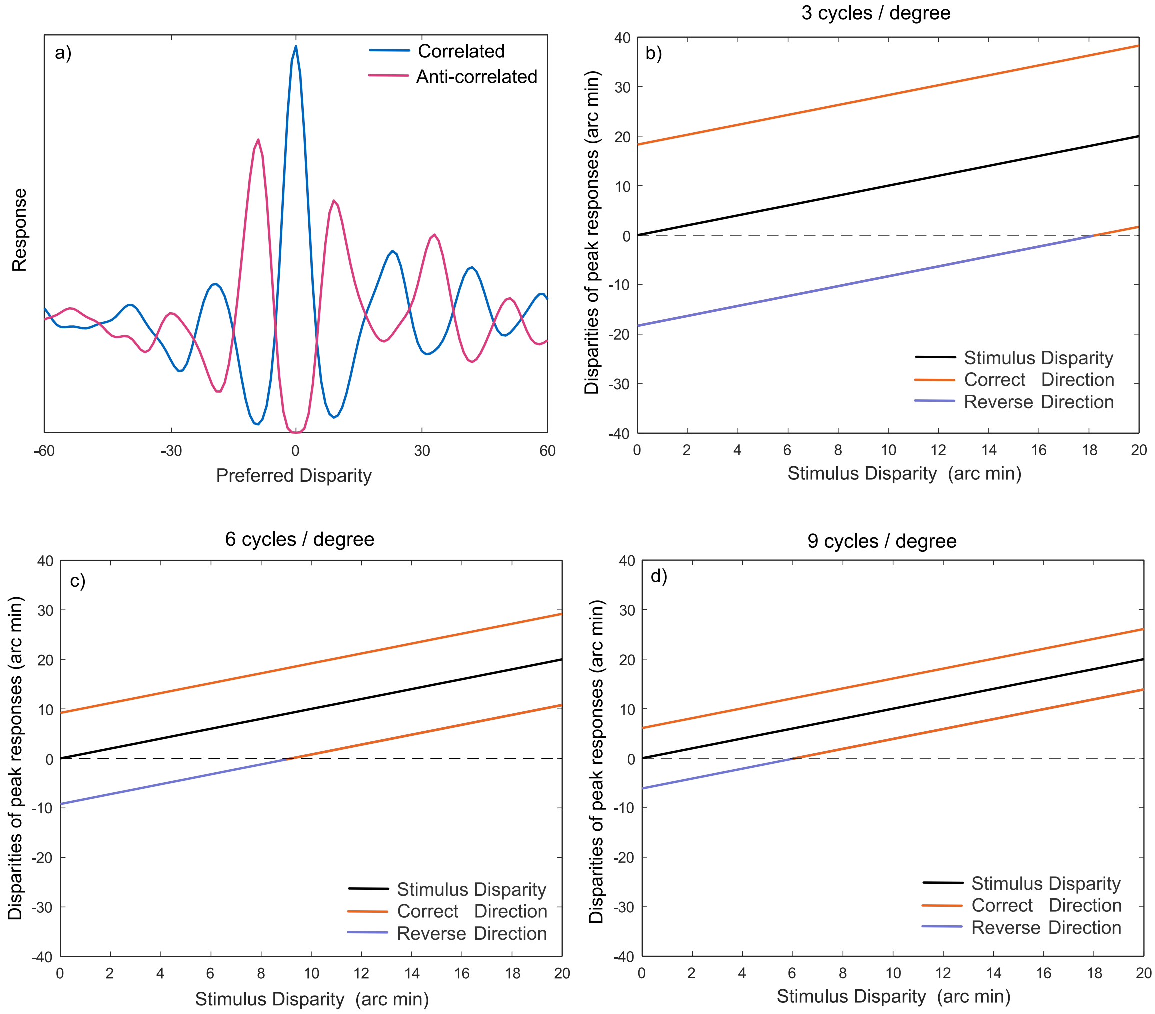


Figure. 6.3 (a) An inverted disparity tuning function in response to an anti-correlated stereogram. The asymmetry evident in this example is due to random sampling variation; the key characteristic prediction of the energy model is the inversion of the function in comparison with the response to a correlated pattern. (b-d) This produces peak responses that are offset from the tuning of the filter for correlated stimuli, which can signal the opposite sign of disparity to that of the neuron's preferred disparity. The sign of this false peak depends on the tuning frequency of the filter and the stimulus disparity. Each plot shows the disparities of the peak responses to anti-correlated stimuli, for a filter that is tuned to the same disparity as the stimulus. Two peaks are plotted, one at a more crossed disparity, and one at a more uncrossed disparity, as indicated in (a). Responses are plotted as a function of disparity, averaged over 1000 Gaussian random noise stimuli. Individual plots show results for three different spatial frequencies. Depending on the combination of the stimulus disparity and the spatial frequency tuning of the filter, the two peaks either both signal forward-depth, or signal conflicting directions of depth.

The perception of depth in ACRDS is further complicated by the existence of both first- and second-order mechanisms in stereoscopic processing. The discussion above considers the disparity information present in the Fourier components of the image, and how they might be combined. However, depth can also be perceived in second-order stimuli, containing informative disparities in variations in contrast, rather than in the underlying texture (Wilcox & Allison, 2009). Evidence for the existence of second-order channels has been provided by both psychophysical experiments, showing that participants can perceive depth in these stimuli, (Edwards et al., 1999, 2000; Hess & Wilcox, 1994; Langley et al., 1998, 1999; McKee et al., 2004; Schor et al., 1998; Wilcox & Allison, 2009; Wilcox & Hess, 1995, 1996, 1997) and physiological experiments, showing disparity-tuned responses to contrast envelopes (Tanaka & Ohzawa, 2006). Tanaka and Ohzawa (2006) accounted for the responses of second-order neurons using a variation of the energy model. This model takes as its input not raw image values, but the outputs of monocular energy filters (Figure 6.4). This monocular energy calculation is followed by a binocular energy calculation, with filters tuned to a much lower spatial frequency.

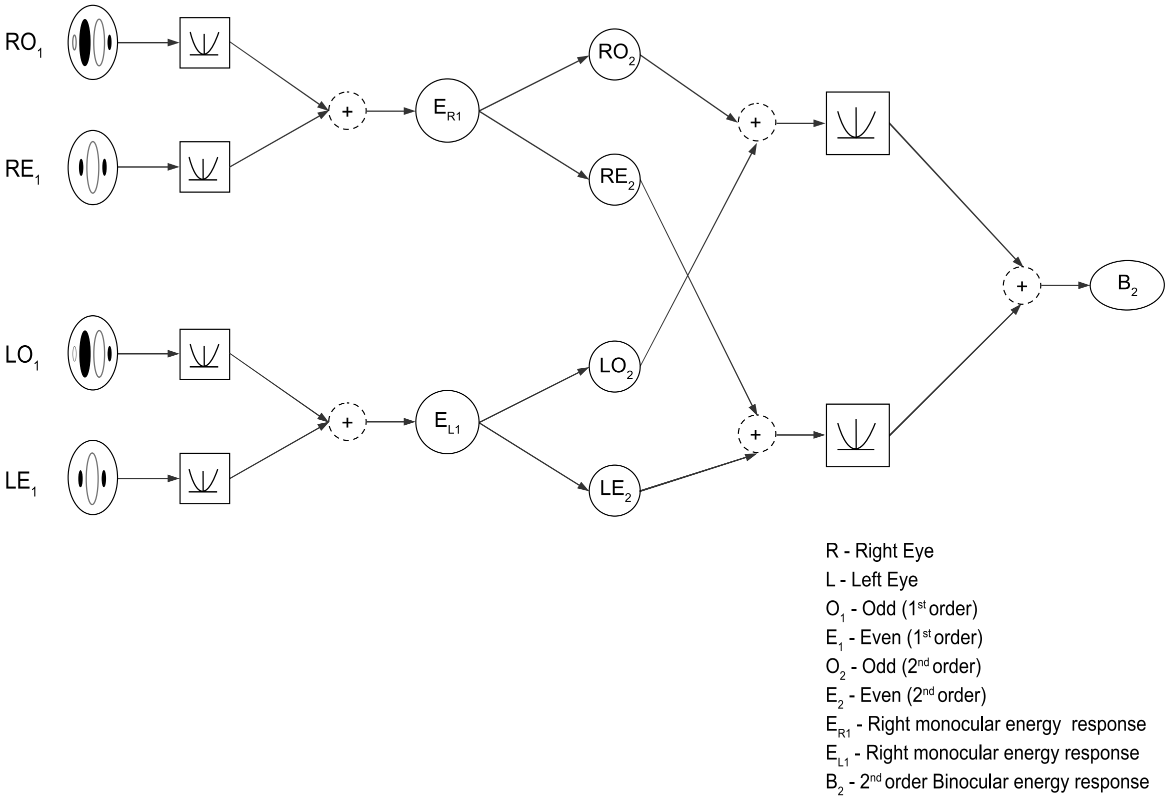
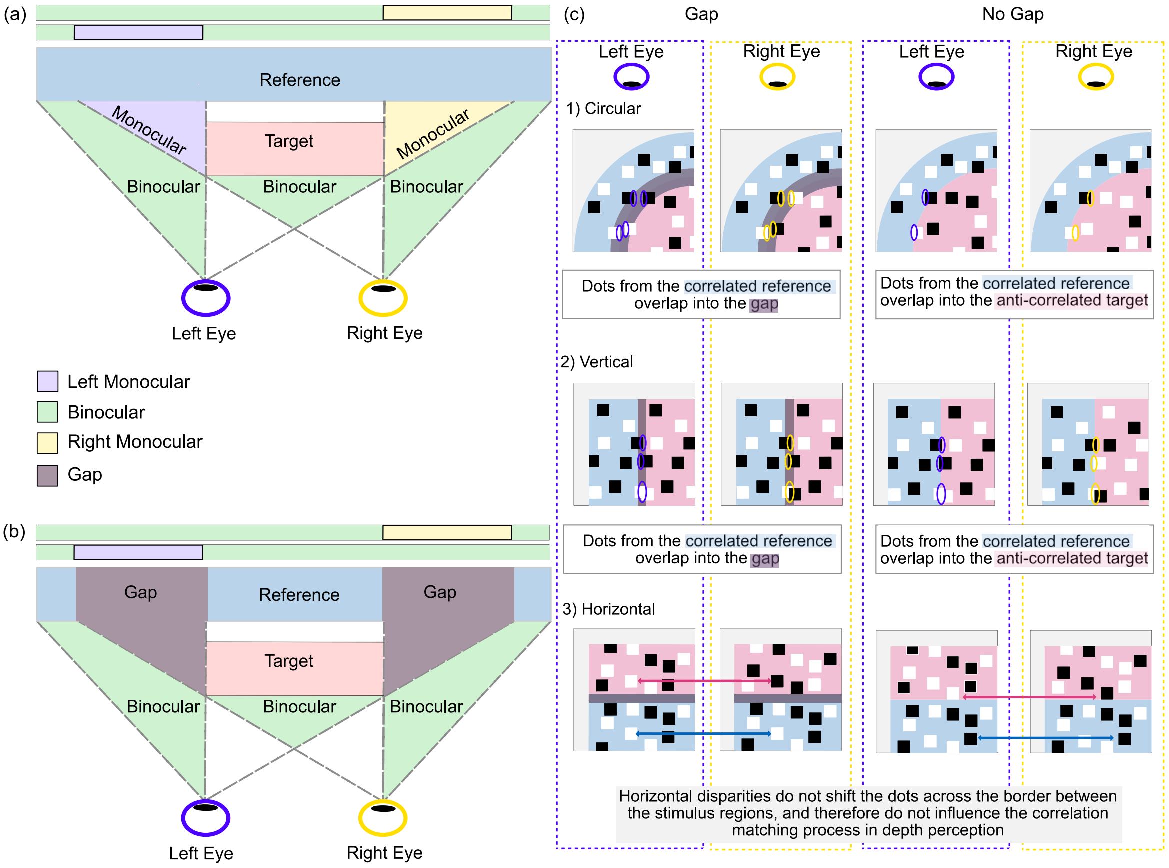


Figure. 6.4 A second-order binocular energy model. The monocular energy responses are calculated separately for each eye. The first-order filtering responses are squared and summed, to produce monocular energy responses (E_{R1} and E_{L1}). This forms the input to the second order filters where the pairs of monocular filters are summed, squared and combined to calculate the second-order binocular energy response.

Wilcox and Hess (1996) showed that depth can be perceived from contrast envelopes even when the underlying random noise patterns are completely uncorrelated, provided that there is an informative disparity in the contrast variation. Sensitivity to depth from second-order stimuli is much poorer than that for first-order stimuli, (Wilcox & Hess, 1997) and allows simple depth judgements to be made, but not the perception of 3D shape (Ziegler & Hess, 1999). Second-order mechanisms will also provide disparity-tuned responses to reversed polarity stimuli, because the rectifying nonlinearity captures the magnitude, but not the contrast polarity, of luminance variations. An important distinction between first- and second-order channels is that the latter will signal depth in the forward, rather than reversed, direction. Cogan et al. (1993) proposed that depth from ACRDS relies on second-order mechanisms. In their experiments, they found forward-depth perception for low density stimuli, but no reliable depth discrimination for high-density ACRDS.

The perception of depth in both CRDS and ACRDS also depends strongly on the presence of features at different disparities in the stimulus, such that the depth of a target can be judged based on its disparity relative to that at other locations. Large changes in absolute disparity, when they are not accompanied by changes in relative disparity, can go unnoticed by participants (Erkelens & Collewijn, 1985a,b; Regan et al., 1986). Sensitivity to depth differences also falls as the spatial separation between target and reference increases (Cottereau et al., 2012) Kamihirata et al. (2015) argued that, due to the relatively weak disparity provided by ACRDS, depth is only perceived when a correlated reference is present, and there is no spatial gap between this and the anti-correlated target. Using stimuli in which the target was a circular region of anti-correlated dots, surrounded by an annulus of correlated dots, they showed that reversed-depth was perceived when there was no-gap, but that this deteriorated when a small-gap was presented. The lack of a gap means that, once a non-zero disparity is incorporated into the stimuli, there will be overlap between the dots in the target and surround, such that both correlated and anti-correlated dots will fall into the receptive fields of neurons aligned with the vertical edges of the stimulus. This creates regions of decorrelation at the edges of the stimuli, on a different side in each eye. Decorrelation of this type occurs naturally through half-occlusion, whereby parts of a stimulus are visible to one eye but not the other (Figure 6.5) (Cook & Gillam, 2004; Gillam & Nakayama, 1999;

Nakayama & Shimojo, 1990; Tsirlin et al., 2010, 2014). Depth is perceived in these stimuli, consistent with this geometric interpretation. It is thus possible that the perception of depth in this case relates to the presence of decorrelation, although it should be noted that this would not explain depth discrimination in other cases (Read & Eagle, 2000) in which there was a horizontal edge.



Dots from the correlated reference overlap into the gap

Dots from the correlated reference overlap into the anti-correlated target

Dots from the correlated reference overlap into the gap

Dots from the correlated reference overlap into the anti-correlated target

Horizontal disparities do not shift the dots across the border between the stimulus regions, and therefore do not influence the correlation matching process in depth perception

Figure. 6.5 (a) Da Vinci stereopsis: when an object (Target) occludes different parts of a background (Reference) from each eye's perspective, there is a difference between the images in the left and right eyes (Nakayama & Shimojo, 1990) However, (b) Including a gap between the Target and the Reference, reduces or eliminates this monocular occlusion. (c) A schematic representation of the process of Da Vinci stereopsis for different types of edges. The different image for each eye can contain matched (correlated) and unmatched (uncorrelated) information. With the exception of a horizontal condition (3), stimuli without a gap between the target and the reference (such as 1 and 2) create a potential area of overlap between correlated and uncorrelated information.

The first purpose of the current study was to understand the contribution of decorrelated edge regions, and the spatial separation between target and surround, on the perception of depth in ACRDS. This was done by using stimuli consisting of (i) an anti-correlated circular target surrounded by a correlated annulus (ii) a vertical edge between correlated and anti-correlated regions and (iii) a horizontal edge between correlated and anti-correlated regions. In the first two stimuli, there are regions containing both correlated and anti-correlated dots, whereas in the horizontal stimuli each region contains exclusively correlated or anti-correlated dots. The second purpose was to understand the contribution of mechanisms sensitive to first- and second-order disparities, and to monocular image regions, to the perception of depth in RDS. We did this firstly by modelling the responses of first- and second-order mechanisms to our stimuli, and secondly by manipulating the presence of decorrelated regions in our psychophysical experiments.

6.2 Psychophysical Experiment

6.2.1 Methods

6.2.2 Participants

10 participants (8 females, mean*(std)* age 24.5*(9.6)*) completed the experiment. All had normal or corrected to normal vision, and stereoacuity of at least 50 arc sec, as measured using the Stereo Optical Butterfly Stereotest. All work was carried out in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki). The study

procedures were approved by the University of Essex Ethics Committee (Application No. JA1601). All participants gave informed written consent and received payment for their participation.

6.2.3 Apparatus

Stimuli were presented on a VIEWPIXX 3D monitor, viewed from a distance of 40 cm. The monitor screen was 52 cm wide and 29 cm tall. The screen resolution was 1920 by 1080 pixels, with a refresh rate of 120 Hz. Each pixel subtended 2.4 arc min. Stereoscopic presentation was achieved using a 3DPixx IR emitter and NVIDIA 3D Vision LCD shutter glasses. The cross-talk between the left and right images, measured using a Minolta LS-110 photometer, was 0.12%. Participants' responses were recorded using the computer keyboard. Stimuli were generated and presented using MATLAB and the Psychophysics Toolbox extensions (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997).

6.2.4 Stimuli

Stimuli in all conditions were random dot stereograms, consisting of 12 arc min square red ($27.0cdm^{-2}$) and black ($0cdm^{-2}$) dots against a red ($13.5cdm^{-2}$) background. Historically red has been used in stereoscopic experiments, as it have been shown to reduce the amount of perceptual crosstalk between the boundaries of the stimuli components (Konrad et al., 2000). Equal numbers of red and black dots were presented, with a total density of 1.12 dots/degree^{-2}. In all cases, stimuli consisted of a correlated reference region, presented with 0 disparity, and a test region which was either correlated or anti-correlated. The dots were in all cases randomly repositioned on each frame. The test region was presented with a disparities of ± 5.5, ± 11 and ± 22 arc min. Stimuli were presented in three conditions: circular, horizontal and vertical (Figure 6.6(a-i)).

6.2.5 Circular

The target was a circular region with a diameter of 4.4 degrees. The reference was a surrounding annulus, with no-gap, a small-gap (19.2 arc min) or a large-gap (38.4 arc min) between target and the inside of the reference region (Figure 6.6 (a-c)). The inner diameter of the reference annulus depended on the size of the gap between the test and reference, and the outer diameter was 1.83 degrees larger than the inner radius. The circular stimuli were presented with the centre of the test region 5.5 degrees below the fixation cross.

6.2.6 Horizontal Edge

The reference and test regions were both a rectangle with a width of 5.5 degrees and a height of 2.75 degrees. Both were centered 5.5 degrees to the right of fixation, with the test presented above and the reference presented below the fixation cross. There were three levels of vertical gap; no-gap, a small-gap (19.2 arc min) or a big-gap (38.4 arc min) between the reference and test regions (Figure 6.6 (d-f))

6.2.7 Vertical Edge

The reference and test regions were both a rectangle with a width of 2.75 degrees and a height of 5.5 degrees. Both were centred 5.5 degrees below the fixation cross, with the test presented to the right and the reference presented to the left. As with the previous conditions, there was a horizontal separation of either no-gap, a small-gap (19.2 arc min) or a big-gap (38.4 arc min) between the reference and test regions (Figure 6.6 (g-i)).

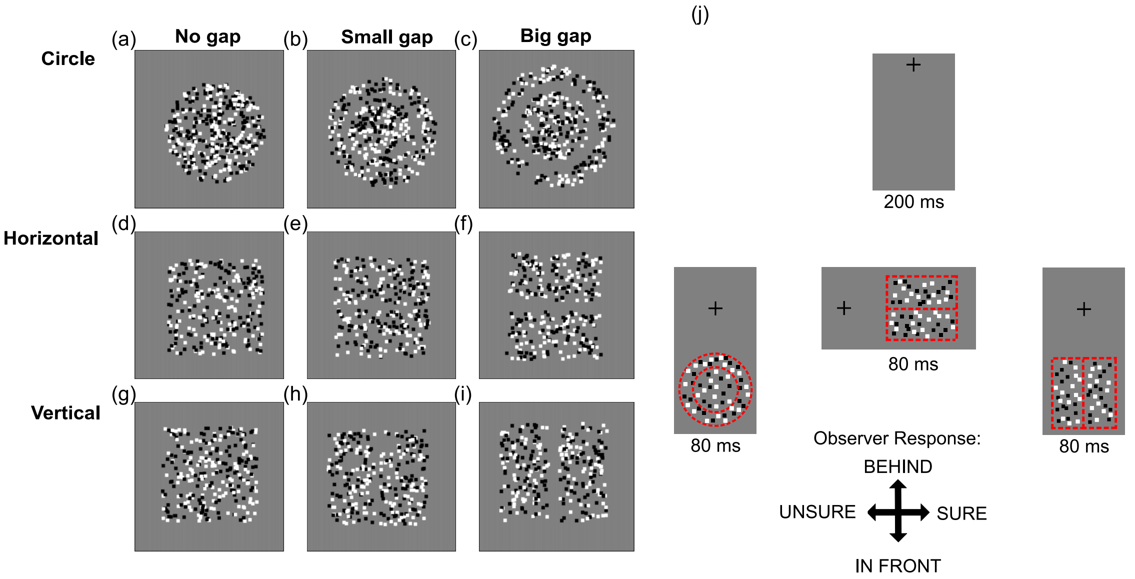


Figure. 6.6 (a) Schematic of stimuli used for each condition. a) For the circular condition, the target was always the inside circle. b) For the horizontal condition the target was always the top rectangle and c) for the vertical condition the target was always the right rectangle. The task in each case was to identify whether the target was behind or in front of the reference. Participants were required to stay fixated on the cross at all times. (b) Participants were required to indicate for each presentation whether the target was in front of or behind the reference, immediately after which a confidence decision (sure or unsure) was required. The procedural schematic shows a border to separate target and reference, and the relative location of the fixation cross for information only.

6.2.8 Procedure

Each trial began with the presentation of a central fixation cross for 200ms, followed by the presentation of a stimulus for 80ms. The fixation cross remained visible throughout each trial block.

After stimulus presentation, the participant was required to make two responses. The first was to indicate whether the target appeared closer (down arrow) or further way (up arrow) than the reference. The second was to indicate whether they felt confident in their response (right arrow) or that they were guessing (left arrow). The next trial began after the two responses had been made (Figure 6.6 (j)).

Each participant completed 18 blocks of trials, for all combinations of the three configurations, three separations, and two correlation conditions. In each block, each of the 6 disparities was presented 20 times. Blocks were presented in a randomised order, and separated over two or more sessions.

6.2.9 Results

6.2.10 Depth Perception

To determine whether participants' judgements shifted from 'far' to 'near' (or vice versa) as disparity shifted from uncrossed to crossed, results were analysed with a generalised linear mixed effects model. In this model, disparity was used as a predictor, with a probit linking function, and random slopes and intercepts across participants were included. A separate model was fit for each combination of shape, correlation and stimulus separation. This model allows us to determine whether, at the population level, forward- or reversed-depth was perceived, while also taking account of variation across participants.

Previous studies have found significant individual differences in depth perception in ACRDS, for example with some participants perceiving reversed-depth but others perceiving no depth (Read & Eagle, 2000). We therefore fit the same generalised linear model, with a probit linking function, to the data for each participant separately. In each case, we then determined whether that participant perceived forward- or negative depth for each stimulus based on whether the slope parameter of the model had a significant positive or negative value. The results are shown in Figure 6.8, which plots the number of participants perceiving forward-depth, reversed-depth, or no significant depth, for each stimulus type.

6.2.11 Circular Stimuli

The proportion of near judgements is plotted as a function of disparity. Results are plotted separately for CRDS and ACRDS and for the three separation distances. Figure 6.7 (a-c) shows the mean results across participants. For CRDS, there was a positive slope for all separations indicating forward-depth. In contrast there was no significant slope in either direction for ACRDS indicating no reliable depth perception (see Table 6.1).

Individual analyses (Figure 6.8) showed that the majority of participants perceived forward-depth for CRDS, while some were not able to reliably report the direction of depth for the larger stimulus separations. For CRDS, in most cases there was no reliable perception of depth. With no separation, 4/10 participants reported reversed-depth,but no reversed-depth was perceived for larger separations, consistent with recent findings (Kamihirata et al., 2015). However, for the largest separation, there was evidence of reliable perceived depth in the forward direction for two participants.

6.2.12 Horizontal Stimuli

Data are presented in Figure 6.7 (d-f) in the same format as for circular stimuli. For CRDS, there as a positive slope for all separations. For ACRDS, there was a significant positive slope for the small separation, consistent with forward, rather than reversed, depth (Table 6.1).

Individual analyses (Figure 6.8) showed that the majority of participants perceived forward-depth for CRDS, although this proportion decreased with increasing separation. For ACRDS, one participant perceived depth in the reverse direction when there was no-gap, and a minority of participants perceived depth in the forward direction.

6.2.13 Vertical Stimuli

Data for the vertical stimuli are presented in Figure 6.7 (g-i). For CRDS, there is a positive slope for all separations. For ACRDS, there was a significant positive slope for the two non-zero gaps, consistent with forward-depth (Table 6.1).

Individual analyses (Figure 6.8) showed very similar results to those found with a horizontal edge. The majority of participants perceived forward-depth for CRDS, and this proportion decreased with increasing separation. For ACRDS, the majority of participants did not reliably perceive depth, but a minority did consistently perceive forward-depth, and there was only one example of reliable reversed-depth.

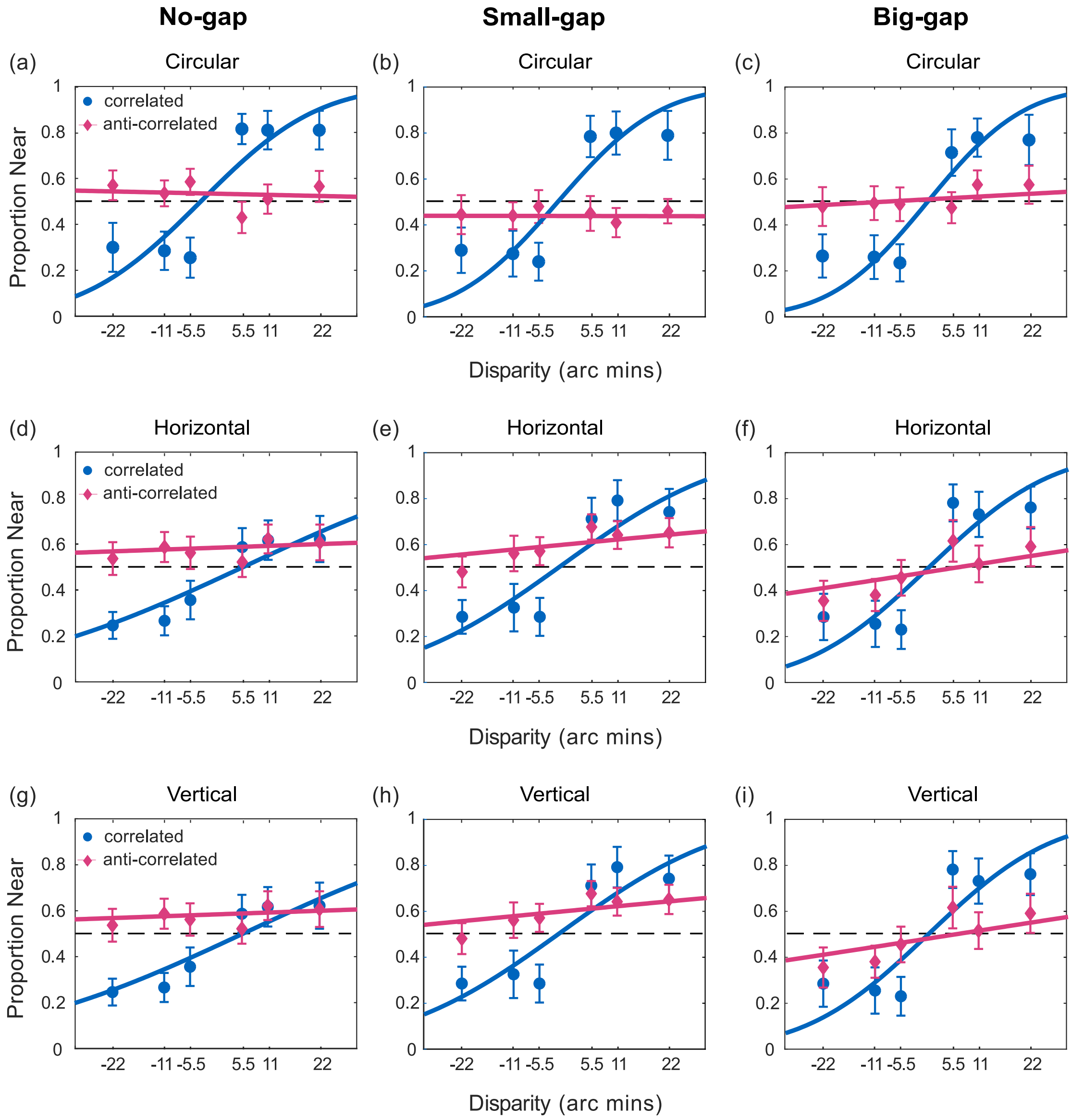


Figure. 6.7 Mean proportion of near responses as a function of disparity for stimuli with no-gap (left), small-gap (middle) and big-gap (right). Circular results are shown in (a-c), Horizontal results in (d-f) and Vertical results in (g-i). Separate plots chart the responses to anti-correlated (pink) and correlated (blue) stimuli. The dashed line for each plot reflects performance at chance, uncrossed and crossed disparities are shown with negative and positive values. Error bars are ±1 standard error of the mean.

Table 6.1 Estimates of effects from the generalised linear mixed effects regression for the proportion of near versus far responses in CRDS and ACRDS stimuli

Shape	**Correlation**	**Gap Size**	β	**SE**	**tStat**	**DF**	**pValue**
Circular	CRDS	0	0.05095	0.018469	2.7586	58	0.0077524*
Circular	CRDS	19.2	0.058554	0.020582	2.845	58	0.0061278*
Circular	CRDS	38.4	0.061829	0.021556	2.8682	58	0.0057471*
Circular	ACRDS	0	-0.0011257	0.0021927	-0.5134	58	0.60962
Circular	ACRDS	19.2	-7.8727e-05	0.0013605	-0.057867	58	0.95405
Circular	ACRDS	38.4	0.0027754	0.0021185	1.3101	58	0.19532
Horizontal	CRDS	0	0.042448	0.015851	2.6779	58	0.0096197*
Horizontal	CRDS	19.2	0.038207	0.015074	2.5347	58	0.013976*
Horizontal	CRDS	38.4	0.039287	0.015882	2.4736	58	0.016324*
Horizontal	ACRDS	0	0.004145	0.0030626	1.3534	58	0.18117
Horizontal	ACRDS	19.2	0.0069563	0.0029483	2.3594	58	0.021692*
Horizontal	ACRDS	38.4	0.0045466	0.0033093	1.3739	58	0.17476
Vertical	CRDS	0	0.023867	0.011142	2.1421	58	0.036395*
Vertical	CRDS	19.2	0.03677	0.014867	2.4733	58	0.016339*
Vertical	CRDS	38.4	0.048554	0.017404	2.7899	58	0.0071229*
Vertical	ACRDS	0	0.0018335	0.0018846	0.97288	58	0.33465
Vertical	ACRDS	19.2	0.0050179	0.0022086	2.2719	58	0.026817*
Vertical	ACRDS	38.4	0.007975	0.0036747	2.1702	58	0.034096*

*indicates significant p value.

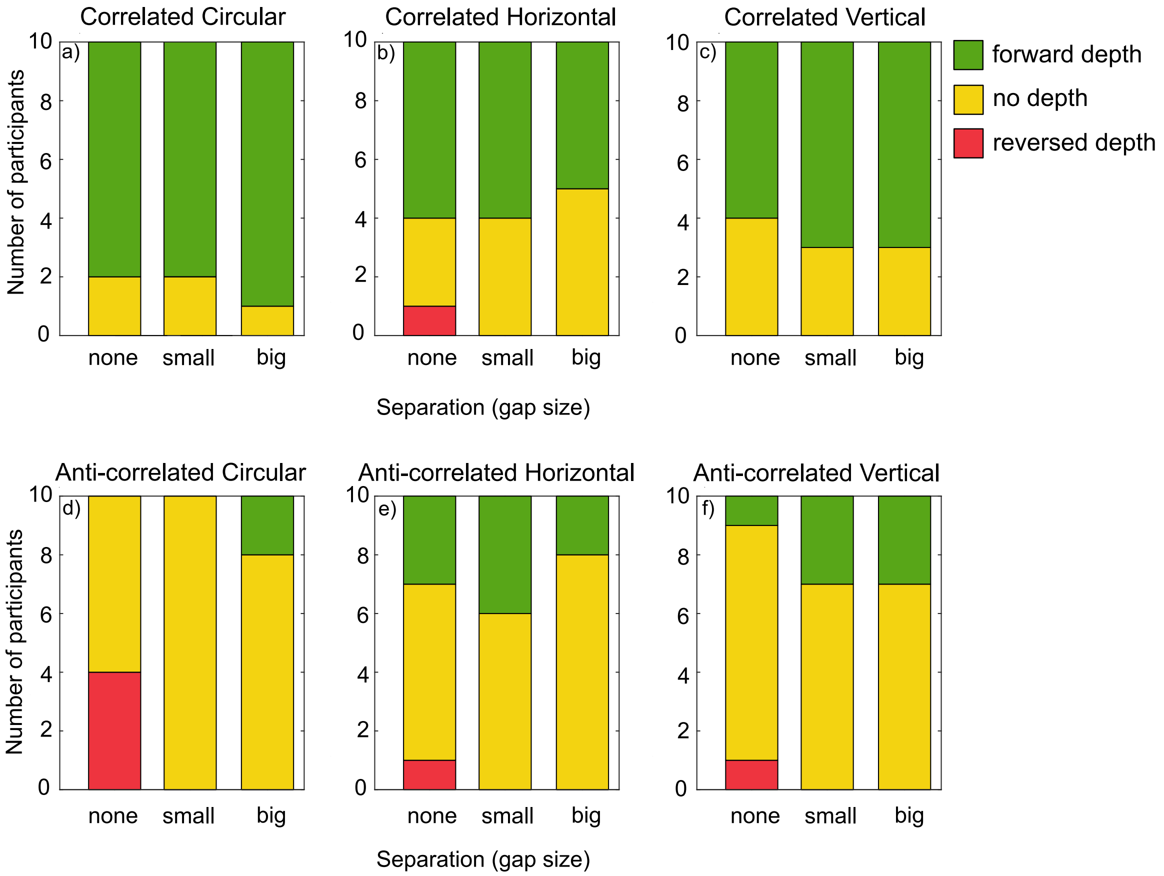


Figure. 6.8 The number of participants reporting forward (green) reversed (red) or no (yellow) depth

6.2.14 Confidence

Mean confidence ratings are plotted as a function Figure 6.10. Results are averaged across signs of disparity since, unlike depth judgements, we do not expect opposite results for near and far stimuli.

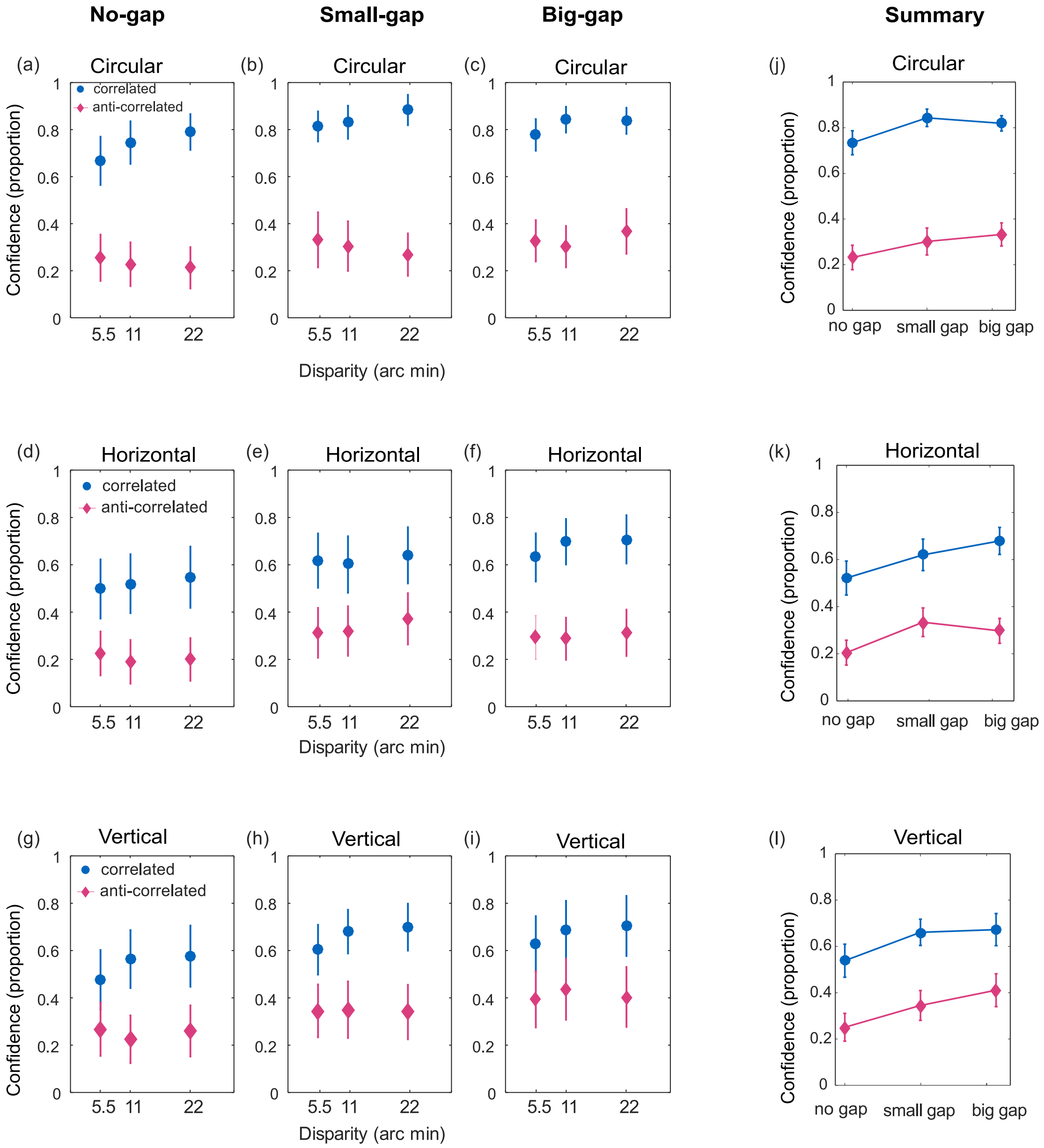


Figure. 6.9 (a-i) Mean confidence (proportion) for CRDS (blue) and ACRDS (pink) are plotted as a function of disparity and gap size and shape. Error bars show ±1 standard error of the mean. (j-l) Summary of confidence in depth detection with each shape on separate plots. Plotted as a function of gap size for CRDS (blue) and ACRDS (pink), an averaged over disparity.

These results were analysed using a linear mixed effects model, with the shape (circular, horizontal or vertical) and correlation (correlated or anti-correlated) as categorical factors, and separation and disparity as linear covariates. A correlation-by-distance interaction term was also included to determine whether separation had a greater effect on confidence for ACRDS than for CRDS. Participant was included as a random factor, with random intercepts and slopes against disparity. The results are summarised in table Table 6.2. Confidence ratings were significantly lower for ACRDS than for CRDS. They were also significantly lower for the horizontal and vertical stimuli than for the circular stimuli. Ratings tended to increase with increasing gap size, but were not affected by disparity. There was no significant distance-by-correlation interaction, meaning that the separation did not affect confidence differently for CRDS and ACRDS. The effects of correlation, shape and distance are summarised along with the mean ratings in Figure 6.9.

Table 6.2 Estimates of effects from the linear mixed effects regression for confidence judgements in CRDS and ACRDS stimuli

Shape	β	**SE**	**tStat**	**DF**	**pValue**	**CI**	**CI**
Circular (intercept)	23.162	3.7038	6.2536	533	8.2323e-10	15.886	30.438
Horizontal	-4.0333	1.0966	-3.678	533	0.00025894*	-6.1876	-1.8791
Vertical	-2.5611	1.0966	-2.3355	533	0.019889*	-4.7153	-0.40689
Correlation	-14.711	2.369	-6.2099	533	1.0675e-09*	-19.365	-10.057
Gap Size	2.5111	0.93709	2.6797	533	0.0075966*	0.67028	4.3519
Disparity	0.083911	0.065264	1.2857	533	0.1991	-0.044296	0.21212
Correlation x Gap	-0.15556	1.0966	-0.14185	533	0.88725	-2.3098	1.9987

*indicates significant p value.

6.2.15 Relationship between Confidence and Performance

For each stimulus, we recorded a near/far depth judgement and an indication of whether or not the participant felt confident in this judgement. Figure 6.10 plots heatmaps of the relationship between confidence and performance, for CRDS and ACRDS, summed over all participants, disparities, separations and stimulus types. Performance was summarised at the number of consistent responses such that, for example, 20 near responses out of 20, and 20 far responses, were both coded as 100% consistent, while a stimulus for which there were 10 near and 10 far responses was coded as 50% consistent. This consistency measure does not depend on whether responses are in the forward or reversed direction, relative to the stimulus disparity. For CRDS, results are clustered in the top-right corner, indicating generally high levels of both consistency and confidence. For ACRDS, confidence was generally low, consistent with a low level of consistency in responses.

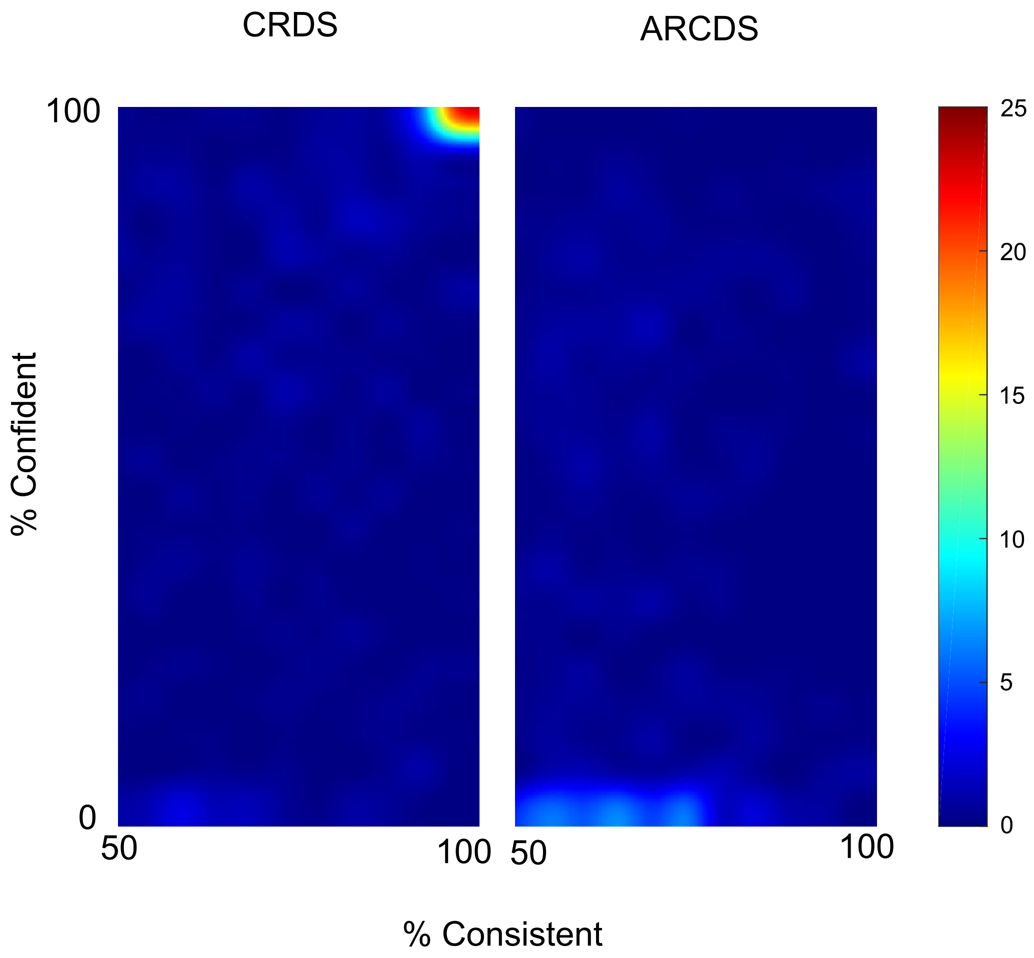


Figure. 6.10 Relationship between consistency of, and confidence in, responses, for CRDS (left) and ACRDS (right). Colour indicates the percentage of responses, for each trial, at each level of consistency and confidence, averaged across all participants and stimuli. High levels of both consistency and confidence were recorded for CRDS, and low levels of both measures for ACRDS.

6.2.16 Discussion

The aim of our psychophysical experiment was to understand how the spatial separation (the size of the gap) and overlapping correlated and anti-correlated elements influence our perception of depth in ACRDS.

Predictably, our results show forward-depth perception for all CRDS stimuli. Depth perception for horizontal and vertical stimuli was not influenced by the size of the gap or type of stimulus. In addition for CRDS, confidence increased with gap size. Confidence was highest for circular stimuli and lower for horizontal and vertical stimuli.

In contrast, for ACRDS confidence in judgements was at or below 50% for all conditions. These are significantly lower than the respective confidence ratings for CRDS. Furthermore, depth perception for ACRDS was variable depending on the type of stimulus.

For circular ACRDS there was, on average, no depth perception for any gap size. However, reversed-depth was seen by some participants in the no-gap condition (Aoki et al., 2017). For the small-gap, there was no depth discrimination shown by any participants, and for the big-gap only two participants reported forward-depth.

For horizontal stimuli, on average there was a tendency towards forward-depth for ACRDS with a small-gap, but no forward or reverse perception of depth for stimuli with no-gap or a big-gap. While the majority of participants reported no depth, a minority reported forward-depth, and one participant reported reversed-depth.

Finally, for vertical stimuli, on average depth was perceived in the forward direction for the small- and large-gap ACRDS, while there was no perception of depth with the no-gap stimuli. This trend reflected the forward-depth perception in a minority of participants. In both the horizontal and vertical edge conditions, perception of forward-depth in CRDS conditions was lower than in the circular conditions. This may be as a result of the surround annulus in the circular condition providing more points of reference (zero disparity), allowing for better judgement of depth for the central target. While it may be of interest to test this edge perception with an inner and outer rectangle, this was not the purpose of our investigation. However, we would predict that the complexity provided by such a target would provide the same results (for CRDS and ACRDS) as the circular conditions. We will discuss this prediction in more detail in the general discussion.

Aoki et al. (2017) found that removing the gap between the target and the surround for ACRDS increased the perception of reversed-depth. However the removal in the gap results in the potential for overlap between the anti-correlated dots in the target and the correlated dots in the surround. This results in decorrelation (an average correlation of zero) at the edges on opposite sides of the target in each eye. This occlusion occurs naturally in the phenomenon of da Vinci stereopsis when the edges of vertically oriented stimuli occlude the surface from the opposite eye (Nakayama & Shimojo, 1990) as illustrated in Figure 6.5 (a-c). If these uncorrelated regions were taken to indicate the presence of half-occlusions, we would predict that depth would be perceived in the reverse direction.

Our results indicate that there was little evidence of reversed-depth for the vertical edges, and one condition (small-gap) showed a significant tendency for depth to be perceived in the forward direction. There was only one participant who reported reversed-depth for the horizontal no-gap condition.

Our results are consistent with the findings of Aoki et al. (2017), where removing the gap did result in reversed-depth for some observers. However, this only occurred for the circular condition, and no reliable perception of reversed-depth was reported for the horizontal or vertical no-gap conditions. This suggests that the complexity of the circular stimulus (particularly when there is no separation between the target and it zero-disparity reference) is more likely responsible for the perception of reversed-depth, than is the removal of the gap.

This study found no robust evidence for reliable reversed-depth perception in any condition, supporting the findings by Hibbard et al. (2014). The anti-correlated circular condition with no-gap produced the highest proportion of reversed-depth perception, consistent with previous results (Aoki et al., 2017; Kamihirata et al., 2015) Furthermore the circular no-gap and small-gap conditions were the only conditions not reporting any forward-depth perception. The trend towards depth perception in the forward direction, while not consistent with the information provided by first-order disparity channels, is predicted by the depth signalled by second-order mechanisms. The conflicting depth signalled by first- and second-order mechanisms for ACRDS is explored in the following section.

6.3 First- and Second-order responses to correlated and anti correlated random dot stereograms

The perception of depth depends on multiple mechanisms, including first- and second- order channels, tuned to a variety of scales and orientations. How each channel contributes to the perception of depth depends on how and where the information they provide is processed. In order to understand the contribution of first- and second-order mechanisms to perceived depth in CRDS and ACRDS, we modelled the responses of these mechanisms to our stimuli. The modelling involved (i) the convolution of the stimulus with the individual left- and right- eye linear filters (ii) the combination of the resulting filter responses to produce the monocular and binocular energy responses described by Equations (6.1) to (6.7) and (iii) identification of the disparity-tuning of the energy units with the largest responses. This was performed using MATLAB code created and described by (Hibbard et al., 2016).

6.3.1 Methods

6.3.2 First- and Second-Order Mechanisms

First-order disparity-sensitive mechanisms were implemented using a standard binocular energy model. The first stage of this model was a quadrature pair of Gabor filters:

$$\begin{aligned} G_{E1} &= \exp\left(-\frac{x^2}{2\sigma_H^2} - \frac{y^2}{2\sigma_V^2}\right) \cdot \cos\left(2\pi f x\right) \\ G_{O1} &= \exp\left(-\frac{x^2}{2\sigma_H^2} - \frac{y^2}{2\sigma_V^2}\right) \cdot \sin\left(2\pi f x\right) \end{aligned} \tag{6.1}$$

where (x, y) is the spatial position, f is the spatial frequency and σ_H and σ_V determine the horizontal and vertical extent of the Gaussian envelope. All filters were tuned to a vertical orientation, and spatial frequencies of 4, 8 and 16 cycles /degree were used. The values of σ_H and σ_V depended on spatial frequency as follows:

$$\begin{aligned} \sigma_H &= \frac{0.39}{f} \\ \sigma_V &= \frac{0.78}{f} \end{aligned} \tag{6.2}$$

The left and right images were both convolved with each filter, to model four responses:

$$\begin{aligned} L_{E1} &= L * G_{E1} \\ L_{O1} &= L * G_{O1} \\ R_{E1} &= R * G_{E1} \\ R_{O1} &= R * G_{O1} \end{aligned} \tag{6.3}$$

Binocular energy was calculated as:

$$B_1 = (L_{E1} + R_{E1})^2 + (L_{O1} + R_{O1})^2 \tag{6.4}$$

The second-order mechanisms included two filtering stages. The first used the same filters as first-order mechanisms. Left and right monocular energy responses were calculated as:

$$\begin{aligned} E_{L1} &= L_{E1}^2 + L_{O1}^2 \\ E_{R1} &= R_{E1}^2 + R_{O1}^2 \end{aligned} \tag{6.5}$$

Second-order filters (G_{E2}, G_{O2}) had the same shape as first-order filters, but a frequency tuning of 0.8 cycles/degree. Monocular responses were then calculated by convolving the left and right energy responses with the second-order filters,

$$\begin{aligned} L_{E2} &= E_{L1} * G_{E2} \\ L_{O2} &= E_{L1} * G_{O2} \\ R_{E2} &= E_{R1} * G_{E2} \\ R_{O2} &= E_{R1} * G_{O2} \end{aligned} \tag{6.6}$$

and combining these to produce the second-order energy response:

$$B_2 = (L_{E2} + R_{E2})^2 + (L_{O2} + R_{O2})^2 \tag{6.7}$$

6.3.3 Stimuli

Firstly, we calculated the first- and second-order binocular energy for responses stimuli with a simple disparity difference. Responses to stimuli with the same random dot pattern as used in the psychophysical experiments were analysed. In each stimulus, the upper half was correlated and presented with zero disparity, and the lower half either correlated or anti-correlated and presented with a disparity of 10 arc min.

Secondly, we calculated responses, as a function of spatial position, to the actual horizontal, vertical and circular stimuli used in the psychophysical experiments.

6.3.4 Procedure

For the first analysis, we calculated the first- and second-order binocular energy responses for 1000 samples of the stimuli. An image was created with a disparity of 0 for all image locations to the left centre (referred to here as the surround), and +10 arc min for all image locations to the right (referred to here as the target). In all cases, the values used to calculate the energy response were taken from the central row of the left image, and from a row in the right image corresponding to a horizontal disparities ranging between ±30 arc min, sampled at intervals of 1 arc min. All convolutions were calculated by multiplication in the Fourier frequency domain.

For the second analysis, we calculated first- and second-order binocular energy responses for 20 simulated trials, using filters tuned to 10 arc min crossed, 0 arc min, and 10 arc min uncrossed disparities, for a stimulus with a 10 arc min crossed disparity. This allowed us to calculate the energy response, as a function of spatial position, from filters tuned to the zero disparity reference, and to the forward and reversed directions of the target disparity.

6.3.5 Results

For the first- and second-order binocular energy responses we used binocular energy units with a phase disparity of 0. Disparity tuning determined by positional shifts in the filters between the left and right eyes. This produces a disparity tuning function that is even symmetric, as shown in the example in Figure 6.11(a & b). In reality, binocular neurons have a combination of phase and position tuning(s) (Prince et al., 2017) and tuning functions will not always be even symmetric. In all cases, however, the binocular energy model predicts that responses will be inverted for anti-correlated stimuli in comparison with responses for correlated stimuli. As well as showing this inversion, responses to anti-correlated stimuli tend to be reduced in magnitude. Alternatively, this can be captured using a more complex model in which a monocular thresholding operation is applied before responses are combined binocularly (Read et al., 2002). Figure 6.11(a & b) display the results of the first analysis, where (a) shows the mean, over 1000 trials, of the first order energy response, as a function of vertical position and horizontal disparity. Results are shown separately for each of the three spatial frequencies and are plotted separately for CRDS and ACRDS. Results were normalised for each frequency by dividing by the maximum of the mean response (Hibbard et al., 2016). The horizontal lines show the target- and surround regions, with a non-zero and zero disparity, for correlated (blue) and anticorrelated (pink) regions. For CRDS, there is a peak in response for all frequencies at the correct disparity, flanked by a minimum in response, the disparity of which depends on the frequency tuning of the filter. This pattern reflects the tuning function shown in Figure 6.1. For the ACRDS, this response is reversed, such that there is a minimum response at the correct disparity. Reversed-depth has been attributed to evidence against the presence of a stimulus at the correct disparity, (Aoki et al., 2017) or to the presence of the flanking false peaks (Read & Eagle, 2000). Figure 6.11(b) shows the mean of the second-order energy response, plotted in the same way as for the first-order response. For CRDS, these results show a very similar pattern to the first-order response, except with much coarser spatial resolution, due to the low-frequency tuning of the second-order filters. For the ACRDS, the inversion of the disparity tuning function is absent. This means that, for ACRDS, second-order filters provide a response that is consistent with the forward disparity, and consistent across the three spatial frequencies of first-order filtering.

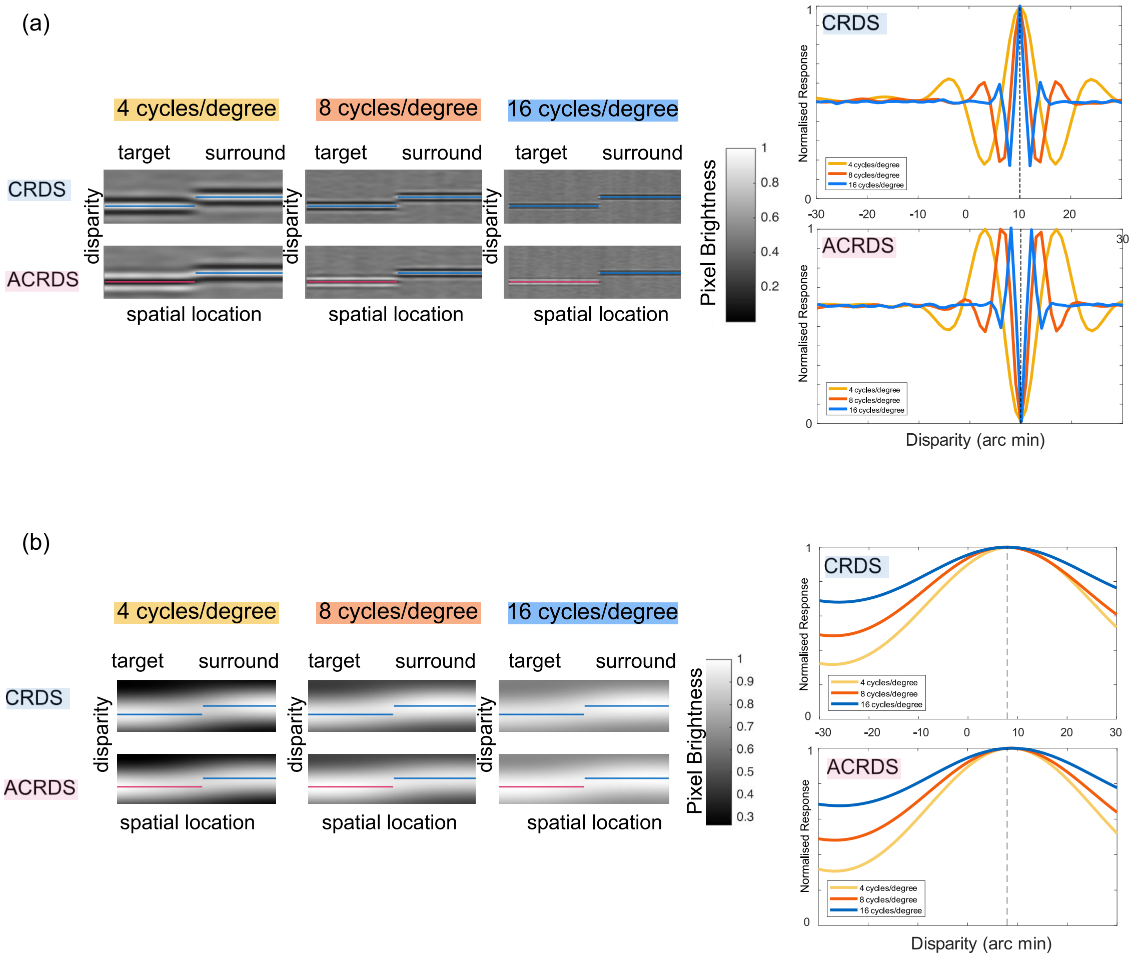


Figure. 6.11 (a) The responses of first-order mechanisms, tuned to 4, 8 and 16 cycles/degree, to CRDS (top row) and ACRDS (bottom row). In each image, horizontal and vertical position represent the spatial location and preferred disparity of the filters, respectively, and pixel brightness the strength of response. The target (correlated or anticorrelated) is on the left of each image and the zero-disparity surround is on the right. The disparities of these are indicated by the horizontal lines (blue for correlated regions, pink for anti-correlated regions). The two graphs on the right show the disparity tuning, averaged over the target regions, for all three frequencies, for CRDS (top) and ACRDS (bottom). (b) The responses of second-order mechanisms, with first-stage filters tuned to 4, 8 and 16 cycles/degree, to CRDS (top row) and ACRDS (bottom row). The second stage filters are in all cases tuned to 0.8 cycles/degree.

The results of the second analysis (Figure 6.12) shows the mean energy responses, over the 20 simulated trials and three spatial frequencies, of the first- and second-order binocular energy responses, at the three different disparities for CRDS and ACRDS for each of the shape and gap conditions.

The strongest response for the correlated zero-disparity reference (surround for circular, bottom for horizontal, and left for vertical) region occurs for the filters tuned to zero disparity for both CRDS and ACRDS in all shape and gap conditions for both first- and second-order mechanisms. For CRDS, the strongest response for the target (centre for circular, top for horizontal, and right for vertical) region signals forward-depth for both first- and second-order mechanisms in every shape by gap condition. In contrast, the strongest responses for the target (centre for circular, top for horizontal, and right for vertical) in ACRDS occurs for the filter tuned to a reversed-depth direction for first-order mechanisms and forward-depth for second-order mechanisms. These results are consistent with previous discussions. Our analysis shows that, for correlated stimuli, both first- and second-order mechanisms signal depth in the forward direction, for all shapes and gap sizes. For anti-correlated stimuli, the first-order channel signals reversed-depth while the second-order channel signals forward-depth.

6.3.6 Discussion

The purpose of the second study was to investigate first- and second-order contributions to the perception of depth. This analysis focused on modelling responses to the stimuli used in this study. The reversal of apparent depth found for some ACRDS has been linked to the inversion of the disparity tuning function of binocular cortical cells. This inversion of disparity tuning does not occur for second-order mechanisms, due to the rectifying non-linearity that occurs

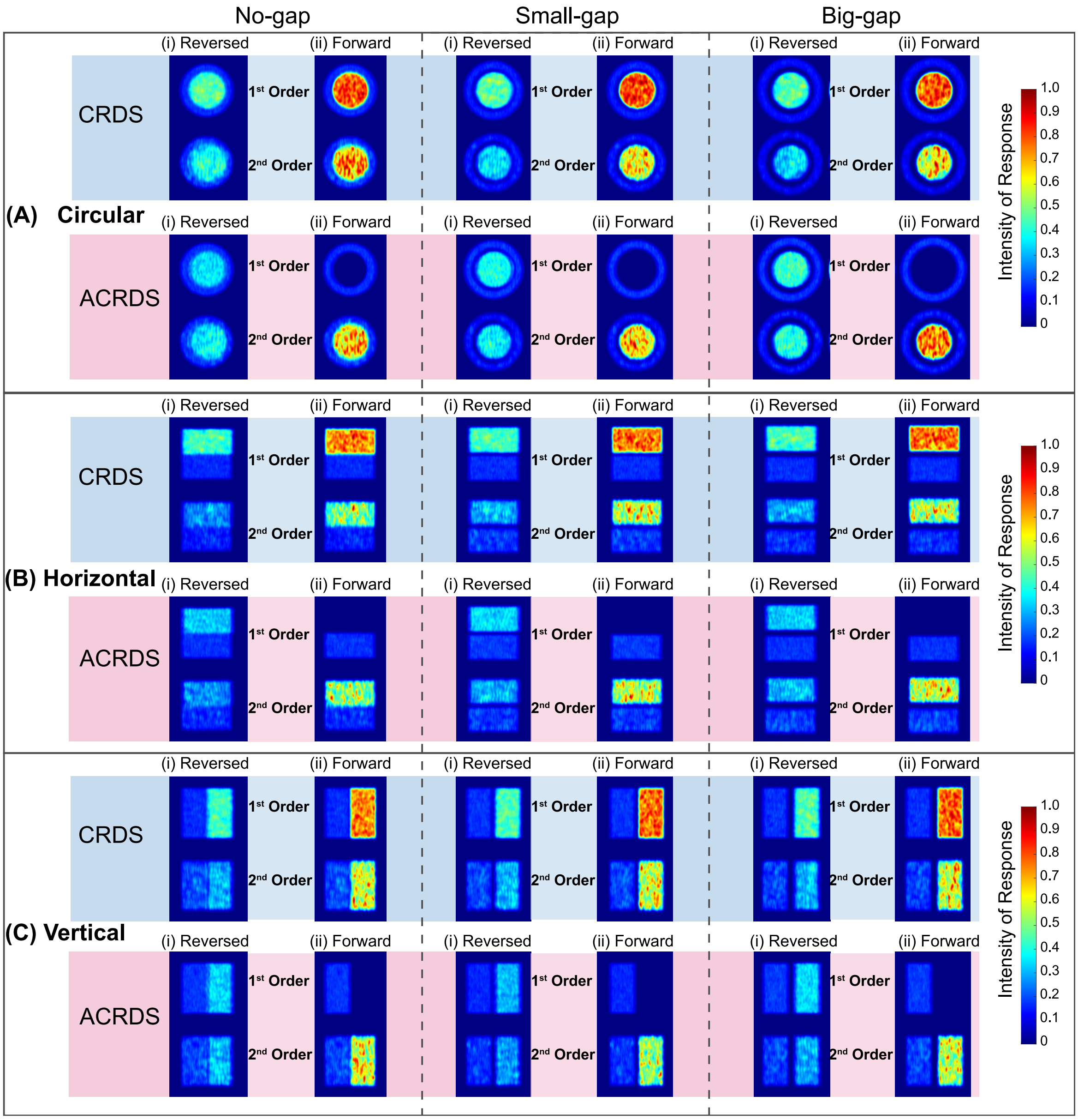


Figure. 6.12 Heatmaps for all nine conditions of the simulation for our specific stimuli. Modelled responses to circular stimuli are on the top (A), horizontal in the centre (B) and vertical at the bottom (C). Each shape has rows for CRDS (top) and ACRDS (bottom), and the gaps in columns, with no-gap (left), small-gap (middle) and big-gap (right). Each heat map contains a response for the target (inner circle, top horizontal and right vertical) for filters tuned to (i) an uncrossed disparity of 10 arc min (reversed), and for the filters tuned to (ii) a crossed disparity of 10 arc min (Forward). Each condition also reports the modelled responses for 1st and 2nd order. The intensity of the response ranges from a zero response (dark blue) to a strong response of 1 (dark red). The first-order channel produces the strongest response in the forward channel for (CRDS), but for (ACRDS) the reverse channel is stronger. In contrast, for the second-order channel a stronger response to the target is evident in the forward direction for both CRDS and ACRDS.

between the first- and second-stage filtering. This means that second-order units indicate depth in the forward direction for both CRDS and ACRDS.

Hibbard et al. (2016) proposed that second-order mechanisms may provide some error correction to the first-order disparity estimations where regions are decorrelated (Hibbard et al., 2016) This is consistent with the perception of forward-depth in ACRDS shown here, despite the lack of a clear binocular cross-correlation peak at the target disparity. While the first- and second-order responses have been shown to be highly correlated in natural stimuli, second-order mechanisms can increase the accuracy of depth perception, acting as a ‘back-up’ mechanism when first-order mechanisms do not provide a reliable signal (Wilcox & Allison, 2009).

The existence of both first- and second-order mechanisms substantially complicates the attempt to link depth perception in ACRDS to neural population responses. It may also account for why there is so much variability across individual participants (Hibbard et al., 2016). As well as providing an impoverished, inconsistent signal across scales in the first-order channels, conflicting estimates can arise between first- and second-order channels.

6.4 General Discussion

Our findings suggest that neither particular spatial configurations of stimuli, nor gap sizes, contributed to robust perception of reversed-depth in ACRDS. As expected, CRDS stimuli were all perceived with depth in the geometrically predicted forward direction. For ACRDS, most participants reported no depth at all, and in some cases there were significant positive slopes consistent with forward-depth perception for the vertical (small-gap & big-gap) and horizontal (small-gap only) conditions. For the anti-correlated circular conditions, overall, there was very little perception of forward-depth for any gap size, with the exception of 2 participants when the gap was at its biggest. For the no-gap condition reversed-depth was reported by 4 participants.

In order to better understand the contributions of the first- and second-order disparity processing stages to the perception of depth, we modelled the responses to the stimuli used in this experiment using a standard binocular energy model, adapted to include second-order filters (Tanabe et al., 2008). Results from the first-order energy response show a peak for CRDS at the correct disparity, as would be expected. For ACRDS, this peak is reversed. The second-order energy response shows a similar peak location as the first-order response for CRDS, but with much broader disparity tuning. However, for the ACRDS there was no inverse peak, as predicted by the rectifying non-linearity occurring prior to the second stage of filtering. This indicates that second-order mechanisms signal forward-depth for both CRDS and ACRDS.

The conflicting signals from the first- and second-order channels may explain the contradictory results between studies (Aoki et al., 2017; Hibbard et al., 2014) and between our conditions, and participants. Wilcox and Allison (2009) proposed that the second-order mechanism provides a back up to the first-order mechanism. When there is a reliable luminance-based disparity signal then the first-order mechanism will be relied upon. However, when there is an ambiguous signal, such as when using anti-correlated or uncorrelated dots, second-order signals are used (Wilcox & Allison, 2009). Strong evidence to support this is found in individuals with strabismus, (McColl et al., 2000) a condition where the eyes are abnormally aligned. Standard clinical tests suggest that individuals with strabismus are stereo-deficient, however McColl et al. (2000) demonstrated that some individuals were still able to perceive depth in their task. Wilcox and Allison (2009) suggest this depth perception was via coarse disparity signals from the second-order mechanisms.

The specific contributions by the first- and second-order mechanisms to depth perception have been described as quantitative and qualitative, respectively (Wilcox & Allison, 2009). Second-order stereopsis provides low-resolution depth estimates over a large range of disparities, where first-order mechanisms provide a high-resolution estimate of depth when the target is matched (correlated) in both eyes. Investigating the contribution of the second-order mechanisms, Ziegler and Hess (1999) used an assortment of stimuli specifically designed to disassociate between linear and nonlinear mechanisms. The task required observers to identify the shape of a target. During testing the display could switch between matched (correlated) and unmatched (uncorrelated) versions of the stimuli. All the observers performed normally and identified shape in matched mode, and as such a staircase converged on their identification threshold. However, when switching to unmatched stimuli, observers could no longer identify the shape, the staircase failed to converge and performance was at chance. Ziegler and Hess (1999) found that while depth perception was demonstrated in all conditions, no reliable perception of shape was reported for any nonlinear (unmatched/uncorrelated) stimuli.

For depth perception, the first-order signal is often the default signal, switching to second-order signals when the first-order is ambiguous (Wilcox & Allison, 2009). This would result in a prediction of forward-depth for all CRDS conditions. However, if the second-order stimulus (for example as seen with no-gap or small-gap) is also ambiguous then potentially there is no useful depth information from either. Consider the heatmaps from our simulations, specifically those for the second-order responses to circular ACRDS (Figure 6.12(a) bottom row). These are the stimuli for which, in our psychophysical experiment, observers reported the lowest perception of forward-depth. Only the big-gap condition had 2 observers report forward-depth reliably, the no-gap and small-gap conditions had no forward-depth perception Figure 6.8. Therefore, these conditions represent the most challenging case for the second-order mechanism, which according to the model also have predicted forward depth perception.

For anti-correlated stimuli where there was no-gap, the horizontal condition had the highest number of observers (3) reporting forward-depth, with none in the no-gap circular, and 1 in the no gap vertical. This was the only no-gap condition where the target and reference did not overlap. This suggests that while the absence of the gap may result in reversed-depth perception in ACRDS for first-order mechanisms, the second-order mechanisms rely on the presence of the gap to reduce the general ambiguity of the stimulus in order to provide the perception of forward-depth. In the case of the anti-correlated circular no-gap condition the complexity of the shape (target surrounded by reference) meant the second-order mechanism was unable to segment the target from the surround. The results of our computational modelling show that the disparity-dependent responses of the second-order channel were sufficient for this to be achieved for both CRDS and ACRDS, and would thus be able to support the perception of depth in the forward direction in both cases. However, a reliable forward-depth signal from this channel did not appear to be available in the ACRDS circle case. The most likely explanation for this is therefore not that the information was not available from the second-order filters, but that the visual system was not able to harness this for the perception of depth-defined shape. It is well-established that the reliable perception of depth depends on relative disparity signals (Cottereau et al., 2012; Erkelens & Collewijn, 1985a,b; Kamihirata et al., 2015; Regan et al., 1986), and in order for this to be achieved extrastriate mechanisms pool disparity signals across space (Anzai et al., 2011; Assee & Qian, 2007; Bredfeldt & Cumming, 2006; Bredfeldt et al., 2009; Orban et al., 2006; Thomas et al., 2002; Umeda et al., 2007; Zhaoping, 2002, 2005). Ziegler and Hess (1999) psychophysical results suggest that the spatial pooling of second-order signals in order to resolve disparity defined shape is severely limited. It is this limit that is most likely account for the inability of observers to segment the anti-correlated circular target from its surround in our experiments.

Another possible explanation for our results may lie in the depth cue combination model proposed by Landy and Johnston (1995). Depth cues will vary throughout a scene, and while each individual cue contributes to the estimated weighted average of depth, these will vary in the quality of information they contain (Landy & Johnston, 1995). Where a cue does not provide a depth estimate, provides an unreliable estimate, or one that contradicts the other cues, it will be down-weighted. Landy et al. suggest a mechanism of combination that statistically estimates depth from weighted averages which depend on the interactions (Landy & Johnston, 1995) and correlations (Hibbard et al., 2016; Oruç et al., 2003) between responses. While their model describes cues in terms of qualities of the stimulus, such as texture, shading and parallax, we have shown that, even within a cue, conflicts might exist, such as between estimates from first- and second-order mechanisms. Therefore it is likely that the average weighted estimate includes the error checking response of the second-order mechanism before interpreting depth. While it is still unknown whether first- and second-order mechanisms co-exist or transition from one to another, (Wilcox & Allison, 2009) our results suggest that when first- and second-order responses are in agreement, depth is perceived normally. Conversely, when the responses from first- and second-order mechanisms conflict, depth would be expected to be perceived at the stage consistent with the direction indicated by the more reliable channel, with the highest decision weight. In the vertical and horizontal stimuli this was the second-order mechanism. However, the spatial complexity of the circular condition suggests that neither first- nor second-order mechanisms are reliable, and therefore the visual system was unable to unambiguously reconcile depth. This would reflect the inability of the second-order channel to reliably signal depth other than in stimuli with very simple spatial variations in surface layout (Ziegler & Hess, 1999). The results of our modelling show that the outputs of second-order filters could in all cases be used to support the perception of forward-depth. The lack of depth perception in ACRDS is consistent with the proposed inability of the visual system to use the available information for the perception of depth-defined shape.

Reversed-depth perception is not routinely found in centrally presented targets (Doi et al., 2011; Hibbard et al., 2014), however our results are consistent with a recent paper (Zhaoping & Ackermann, 2017) that found reversed perception of ACRDS in the periphery. Zhaoping and Ackermann (2017) account for their findings through reduced (or absent) feedback connections in the periphery, resulting in a reduction of verification for stimulus features (see Zhaoping and Ackermann (2017) for a detailed discussion of the feedforward-feedback-verify-weight network model). An important observation to make is that the stimuli used in our study were neither strictly centrally nor peripherally presented, but ‘mid-way’ between the two positions used by Zhaoping and Ackerman. Since second-order channels are also strongest in central fields, and weaker or absent in the periphery, this may also account for these findings. Therefore, it is also likely that the reversed perception in this condition (circular no-gap) for some participants, was a result of the combination of the part-peripherally located presentation and the complexity of the stimulus.

Since spatially separating the reference and the surround has been found to reduce stereoacuity (Aoki et al., 2017; Cottereau et al., 2012), it has been suggested that the spatial gap may be responsible for the lack of robust reversed-depth perception in ACRDS. Consistent with Kamihirata et al. (2015) we found reversed-depth when there was no-gap, which diminished as the gap size increased. However, neither the horizontal nor vertical conditions showed this trend. This implies that the absence of the gap is not exclusively responsible for promoting the perception of reversed-depth. The relatively complicated shape of the circular target and its annulus when there is no-gap and thus the opportunity for overlap also plays a role in reversed-depth, through reducing the effectiveness of the second-order channel to signal forward-depth. Cumming et al. (1998) found that reversed-depth perception from ACRDS reduced as a function of increased dot density. Perception of reversed-depth has been found at low dot densities (between 1% and 11%) (Cogan et al., 1993; Cumming et al., 1998). At high dot densities it is more likely that there is overlap between the dots in each eye, increasing the chance of an ambiguous signal.

A clear finding arising from ACRDS studies (Aoki et al., 2017; Cumming et al., 1998; Hibbard et al., 2014) is that there is considerable variability between individuals' reported perception of depth. Recent findings for robust reverse perception of ACRDS in the periphery (Zhaoping & Ackermann, 2017) suggest that ACRDS still has an important role to play in allowing us to understand how the visual system learns to process depth. Cumming et al. (1998) attempted to improve depth perception for ACRDS using training with feedback, but were unsuccessful. They used trial by trial feedback with over 10 000 presentations of ACRDS, yet performance did not improve. Recent advances in perceptual learning research indicates that crucial factors in learning are individual confidence and task difficulty (Jeter et al., 2009; Lin et al., 2017; Talluri et al., 2015). Our results show a very clear difference between participants' confidence in their judgements for CRDS and ACRDS. Perceptual learning to improve depth perception in anti-correlated stimuli may benefit from these advancements by focusing on improving participant confidence.

6.4.1 Conclusion

How the visual system combines monocular sensory information into a binocular percept is an important question in vision research. Based on the binocular energy response from first-order mechanisms, it has been proposed that reversed-depth perception should occur when presenting an ACRDS target with a CRDS surround. However, these predictions have only taken into account the first-order responses, ignoring the squaring non-linearity of the error checking, (Hibbard et al., 2016) 'back up' function (Wilcox & Allison, 2009) of the second-order mechanism. Our results and analyses underline the need to take account of the interactions between multiple multiple binocular mechanisms when predicting the perception of depth. The sometimes conflicting signals available from ACRDS, containing disparities which are not consistent with geometrical considerations, mean that such predictions are by no means straightforward. Our results are consistent with an increasing number of studies that show second-order summation occurs as an independent process taking place after first-order processing, in motion, (Edwards & Metcalf, 2010; Edwards & Nishida, 2004) perceived contrast (Baker et al., 2012) and binocular phase (Zhou et al., 2014) and stereopsis (Wilcox & Hess, 1996) Finally, while depth perception for ACRDS was mostly ambiguous, the positive slopes found in some conditions indicate a trend for perception in the forward direction, rather than reversed-depth. Furthermore, confidence for ACRDS was poor in every condition, and increasing participant confidence may be the key to improving reversed-depth perception.

Chapter 7

Summary and Conclusion

7.1 Summary of Aims

Perceptual learning is a phenomenon that describes the improvement in specific perceptual tasks as a result of practice (Fahle, 2002). The topic remains one of the most studied phenomena in perception (Maniglia & Seitz, 2018). Understanding the mechanisms of visual perceptual learning is one of the primary objectives for perceptual learning researchers, and findings of perceptual learning are robust and yet seemingly inconsistent. There are many factors that may influence visual perception and learning, and these were covered comprehensively in Chapter 1, are discussed briefly below, and formed the basis for our investigations. One aspect of visual perception that is usually not taken into account is the individual variability in participants' responses (Mollon et al., 2017).

In Chapter 2 we provide a discussion on the difficulties and complexities of analysing heterogeneous data. We describe the benefits of using mixed effects regression, that simultaneously fits a population level psychometric function and provides a statistical analysis of the data. Modelling the data using mixed effects methods allows data that are heterogeneous, such as those with unreliable performance or missing data points, to be included in the analysis. Embracing this variation is argued to be a far more powerful and representative analysis of the broader population (Agresti, 2013; Baldassarre et al., 2012; Baldassarree et al., 2016; Bolker et al., 2009; Halpern et al., 1999; Kosslyn et al., 2002; Mollon et al., 2017; Moscatelli et al., 2012; Nefs et al., 2010). The statistical methods in every experiment in this thesis use these methods and thus a dedicated chapter has been used to fully describe the benefits of the model and directly compares and equates it to the methods used as standard in psychophysical analysis. This statistical method was used to evaluate the difference in populations where a difference in performance was expected.

In Chapter 3 we used an established set of stimuli developed by Polat and Sagi (1993) to evaluate the potential differences between the lateral interactions in neural responses between individuals who suffer migraine and a control population. There is physiological (Bridge et al., 2015) and psychophysical (O'Hare & Hibbard, 2016) evidence that suggest an impairment in the lateral interactions of individuals who suffer migraine. Individuals with migraine have been evidenced to exhibit many sensory processing deficits in tasks including sensitivity to spatial contrast (Benedek et al., 2002; McKendrick & Sampson, 2009), colour and orientation (Palmer et al., 2000; Shepherd, 2000), global motion and form (Antal et al., 2005; Ditchfield et al., 2006; McKendrick et al., 2001). Since low-level visual perception is highly dependent on the lateral interactions between neurons (see Figure 1.3 in Chapter 1) our aims were to investigate if the excitatory and inhibitory lateral interactions may differ in people who have migraine-with-aura. We predicted that should the lateral interactions differ between the groups, then the response patterns showing inhibition and facilitation (based on the size of separation between the target and a pair of colinear flankers) would differ between groups too.

The primary motivation for this thesis was the potential of perceptual learning to aide individuals with a visual impairment. Historically, improvement in visual perception was believed only to be effective when applied in early childhood (Levi, 2005; Levi et al., 1997). However, more recently evidence is accumulating suggesting that perceptual training can improve perception for individuals including adults with a range of impairments such as amblyopia (Astle et al., 2011a,b; Hess et al., 2003; Levi & Li, 2009; Levi & Polat, 1996; Levi et al., 1997; Polat et al., 2004), myopia and presbyopia (Camilleri et al., 2014b; Polat, 2009), nystagmus (Campana & Maniglia, 2015), steropsis (Ding & Levi, 2011), cortical blindness (Das et al., 2014; Huxlin et al., 2009; Kasten & Sabel, 1995; Nelles et al., 2001; Sahraie et al., 2006; Trevethan et al., 2012) and macular degeneration (Maniglia et al., 2016; Plank et al., 2014). However perceptual learning as a tool for rehabilitation is limited due to the highly specific nature of improvement, which is restricted to the features of the trained stimulus or task (Dosher & Lu, 2017). Based on the specificity to low-level features, early perceptual learning researchers argued for a low-level neural location for the learning mechanism (Gilbert et al., 2001; Ren et al., 2016; Schoups et al., 2001). More recently there is growing evidence for a higher level (non sensory area) mechanism, which has been modelled using a computational neural network approach to explain learning as a process of reweighting perceptual decision making units (Dosher et al., 2013; Dosher & Lu, 1998, 2009; Herzog & Fahle, 1999; Petrov et al., 2006b; Poggio et al., 1991). Behavioural studies of perceptual learning are unable to directly distinguish between improvements that result from a change in sensory encoding, and those that result from a reweighting in a perceptual decision unit (Petrov et al., 2005). However, using a perceptual learning experiments where the sensory representations are the same, but which targets multiple different decision stages, may provide a method to dissociate the learning mechanisms. This was our objective for the experiment in Chapter 4 where we trained observers on a high-level task and looked for improvements in both high and low level areas. However, prior to this we queried if feedback in terms of external trial-by-trial responses to indicate accuracy was a requirement for learning. To address that we conducted a pilot study first which identified that in our specific task, trial-by-trial feedback was a requirement.

In Chapter 5 we explore the characteristics of feedback in perceptual learning. Since natural selection would favour those organisms who could sift through conflicting information (Nessler et al., 2008), making use of performance feedback would 'quickly evaluate ones' learning strategy (Shibata et al., 2009). Feedback can improve the efficiency of perceptual learning but not reliably so (Herzog & Fahle, 1998; Seitz et al., 2006) and empirical evidence as to its necessity is mixed. Learning has been found to occur with and without feedback (Petrov et al., 2006a). Furthermore, there is evidence that interleaving easy and difficult trials during training can sometimes elicit improvement without the need for external feedback (Fahle & Edelman, 1993; Liu et al., 2012a). The aim of this chapter was to resolve some of the conflicting findings withing perceptual learning research. We considered a number of factors that are known to influence learning. Our first factor was the level of processing of the stimulus used for training, local and global tasks are known to be processed at different cortical locations (Zeki, 1974). Global tasks require integration and segmentation which is a cognitively more demanding task, and feedback may be more beneficial in these conditions (Burr & Thompson, 2011; Gilbert et al., 2001). Secondly we queried if the neural loci of the learning mechanism may differ based on the type of global task. To that effect we used two different global tasks (form and motion). The third factor was the presence (or absence) of feedback, and the final factor was the method of presenting the training stimuli. Researchers often use adaptive staircases to present their stimuli, however there has been some suggestion that this may bootstrap learning, and may account for why some studies do or do not find learning (Seitz et al., 2006). In this large scale multi-factor experiment we trained half of the groups without feedback (during the training phase) and tested all groups each day before and after training. We predicted that we should secure robust improvement with and without feedback for local tasks. If learning differs at the global level this would be evidenced in the no-feedback conditions of both (or either) motion and form tasks. Finally, we predicted that when training without feedback, groups trained with MOCS were less likely to show improvement.

The study conducted in Chapter 6 was the only task for which we purposely collected a measure of confidence. The motivation for doing so is that the perception of stereoscopic depth, particularly in anticorrelated stimuli containing conflicting information, is a notoriously difficult task for the general experimental participant (Sagi, 2011). Some of the earliest perceptual training delivered experimentally was to teach inexperienced observers how to fuse each eye's image into a binocular percept (Julesz, 1971). To investigate stereoscopic depth in a lab, researchers use random-dot stereograms (Figure 6.2 (a - b), which present the same image to each eye, introducing a shift in one of the images forces the perception of depth. However, it is predicted by the first-order energy response of a binocular neuron Figure 6.1 (b), that if you present the same image to each eye, but reverse the luminance of the matched pairs, it is possible to see reversed depth (i.e. see a target closer to a reference point when it is actually further away, or vice-versa). A recently published study (Aoki et al., 2017) found reversed depth when removing the gap that separates correlated and anti-correlated areas of the stimulus that reversed depth perception was robust. We tested the perception of reversed depth in anti-correlated random-dot stereograms using three spatial gaps and three types of edge to investigate how this may facilitate or inhibit the perception of reversed depth.

7.2 Summary of Empirical Results

Contrast sensitivity and lateral interactions in migraine

Chapter 3: Individuals with migraine show differences in visual perception compared to control groups. It has been suggested that differences in lateral interactions between neurons might account for some of these differences. This study sought to further establish the strength and spatial extent of excitatory and inhibitory interactions in migraine-with-aura using a classic lateral masking task. Observers indicated which of two intervals contained a centrally presented, vertical Gabor target of varying contrast. In separate blocks of trials, the target was presented alone or was flanked by two additional collinear, high contrast Gabors.

Flanker distances varied between 1 and 12 wavelengths of the Gabor stimuli. Overall, contrast thresholds for the migraine group were lower than those in the control group. There was no difference in the degree of lateral interaction in the migraine group. These results are consistent with previous work showing enhanced contrast sensitivity in migraine-with-aura for small, rapidly presented targets, and they suggest that impaired performance in global perceptual tasks in migraine may be attributed to difficulties in segmenting relevant from irrelevant features, rather than altered local mechanisms.

Spatial frequency tuning of perceptual learning and transfer in global motion

Chapter 4: We investigated a) if feedback was required for learning when using an equivalent noise global motion coherence task, and b) the transfer across spatial frequency of training on a global motion coherence task, and the transfer of this training to a measure of contrast sensitivity. For our first experiment two groups, with and without feedback, trained for ten days on a broadband global motion coherence task. Results indicated that feedback was a requirement for learning. For the second experiment training consisted of five days of direction discrimination on one of three global motion tasks (broadband, low or high frequency random-dot Gabors), with trial-by-trial auditory feedback. A pre- and post-training assessment was also conducted, consisting of all three types of global motion stimuli (without feedback) and high and low spatial frequency contrast sensitivity. We predicted that if learning and transfer are cortically localised, then transfer would show specificity to the area processing the task (global motion). In this case, we would predict a broad transfer between spatial frequency conditions of global motion only. However, if transfer occurred as a result of backward generalisation, a more selective transfer would occur matching the low-pass broadband tuning of the area processing global motion. Our training paradigm was successful at eliciting improvement in the trained tasks over the five days. However, post-training transfer to trained or untrained tasks was only reported for the low spatial frequency trained group. This group exhibited increased sensitivity to low spatial frequency contrast, and an improvement for the broadband global motion condition. Our findings suggest that the feedback projections from global to local stages of processing play a role in transfer.

The effect of feedback and stimulus presentation protocol on perceptual learning for local and global motion and form

Chapter 5: The necessity for external reinforcement, such as response feedback, during perceptual training is a topic of much debate. To investigate this in detail we performed a large multi-level experiment, varying the type of training task (Motion or Form), the level of processing (Local or Global), presence of feedback (With or Without) and finally the method of stimulus presentation (Adaptive staircase or Method of constant). 140 participants were assigned to one of ten groups. Detection thresholds were measured daily before and after training. A 75% detection threshold was calculated for every test session which presented No Feedback to any condition, and the threshold from the daily pre-test was used to determine that day's testing levels (65% for difficult trials and 85% for easy trials) as defined by Liu et al. (2012a). The results of our study showed that learning was robust across motion and form tasks, and their processing levels whether local or global. Significant learning occurred for both the feedback and no feedback groups when easy trials were presented, whether using an adaptive staircase or method of constant stimuli. We replicate the findings of Liu et al. (2012a) and extend them to global and motion stimuli. Some differences were identified between the MOCS and QUEST trained results in that at lower stimulus intensities QUEST methods elicited better performance (when there was No Feedback). However at higher stimulus intensities MOCS trained groups performed the same or better than QUEST trained groups. To improve using MOCS at low or sub-threshold coherences, methods that included either a staircase or trial-by-trial feedback was a requirement to obtain learning in this global motion task. This suggests that when there is a high-level of external noise in the stimulus, the internal signal provided by the presence of easy exemplars, at set thresholds facilitated learning when trained on QUEST, but not when trained with randomly interleaved intensities using MOCS. Finally there was some evidence that for local tasks internal reinforcement was more effective that external reinforcement and obtaining learning. In conclusion, interleaving high and low accuracy trials both local and global orientation tasks leads to perceptual learning independent of the external feedback.

First- and second-order contributions to depth perception in anti-correlated random dot stereograms

Chapter 6: The binocular energy model of neural responses predicts that depth from binocular disparity might be perceived in the reversed direction when the contrast of dots presented to one eye is reversed. While reversed-depth has been found using anti-correlated random-dot stereogram (ACRDS) the findings are inconsistent across studies. The mixed findings may be accounted for by the presence of a gap between the target and surround, or as a result of overlap of dots around the vertical edges of the stimuli. To test this, we assessed whether (1) the gap size (0, 19.2 or 38.4 arc min) (2) the correlation of dots or (3) the border orientation (circular target, or horizontal or vertical edge) affected the perception of depth. Reversed-depth from ACRDS (circular no-gap condition) was seen by a minority of participants, but this effect reduced as the gap size increased. Depth was mostly perceived in the correct direction for ACRDS edge stimuli, with the effect increasing with the gap size. The inconsistency across conditions can be accounted for by the relative reliability of first- and second-order depth detection mechanisms, and the coarse spatial resolution of the latter.

7.3 General Discussion

We showed a robust improvement for perceptual learning with or without feedback, when easy and difficult trials are interleaved. Interestingly, we did not find that the learning was negatively impacted by the presentation methods. MOCS did not provide improvements when the task was difficult (global level task) and the stimulus was ambiguous as a result of noise (lower levels of coherence) unless feedback was provided. However, as signal increased the need for external feedback reduced, and those trained with or without feedback achieved similar results. These findings do suggest that when difficulty is high, staircase methods provide a signal of internal reinforcement (when there is no external feedback) that is not provided by the MOCS training, even when easy and difficult trials are presented. When using staircase methods that track a defined threshold groups, improvement was seen in sub-threshold stimuli whereas with MOCS it was not. Conversely at the upper asymptote where signal intensity was higher, performance was better for MOCS trained groups. One of the initial motivations for conducting the study in Chapter 5 was to resolve some of the ambiguity between studies that either did (Liu et al., 2012a) or did not obtain learning in the absence of feedback (Asher et al., 2018b; Seitz et al., 2006) when interleaving easy and difficult trials. The findings from our experiments do not directly resolve all of these differences.

In Chapter 4 we used a global motion training task and, when no feedback was present, there was no learning, secondly when feedback was presented during training, performance returned to baseline (for some conditions) when feedback was removed. One difference in experiments is the way this study added external noise to the global motion task. Typically, global tasks require observers to integrate and segregate information (Burr & Thompson, 2011; Nishida, 2011). In this way, global motion coherence is studied using random dot kinematograms, requiring observers to make a direction judgement from a stimulus comprised of a pattern of moving dots. A typical stimulus, such as the stimulus used in this study, contains a proportion of signal dots moving in one direction and noise dots moving in random directions (Ball & Sekuler, 1982, 1987; Vaina et al., 1995) (shown again in Figure 7.1 (a)). Difficulty is increased by reducing the ratio of signal-to-noise dots; the more noise dots the lower the coherence, and the more difficult the task. In order to perceive a coherent global motion, observers need to integrate and segregate the motion signals over space and time (Burr & Thompson, 2011). However another method of investigating motion coherence

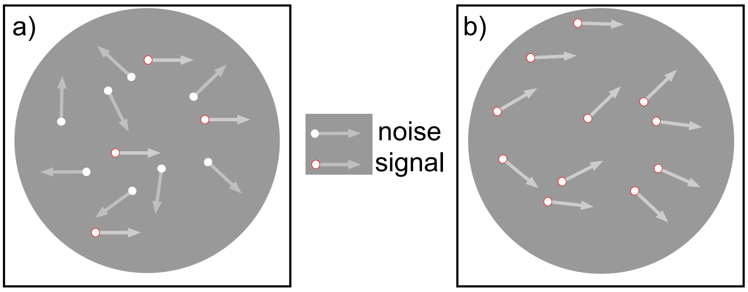


Figure. 7.1 Schematic of global motion stimuli where a) is a typical stimulus comprised of signal dots and noise dots (approximately 30% coherence). b) An equivalent noise stimulus where all dots contribute equally to the signal and the noise (drawn from a distribution of approximately 90° out of a possible 360°). In each case the correct answer would be that the dots are moving towards the right.

is to use an ‘equivalent noise’ paradigm (Dakin & Bex, 2001; Dakin et al., 2005b; Levi et al., 2015; Tibber et al., 2014; Williams & Sekuler, 1984). Rather than having distinct populations of signal and noise dots, all dots contribute both to the signal and the noise by drawing the direction of motion for each dot from a random distribution. Dots move the same distance between each frame and the direction travelled is independent of the directions of the other dots (Williams & Sekuler, 1984). Difficulty is increased by manipulating the standard deviation of the distribution of directions presented, thus each dot contributes to the signal (Tibber et al., 2014) see Figure 7.1 (b).

It has been predicted that populations displaying a deficit in detecting motion coherence, such as those with migraine (Shepherd, 2006; Tibber et al., 2014), amblyopia (Mansouri & Hess, 2006) cortical blindness or in ageing populations (Ostwald et al., 2008) may have a reduced ability to segment out external noise (Tibber et al., 2014). Using an equivalent noise task means that the stimulus elements (much like in a local task) all contribute to the trajectory of the signal, and they do not contain overt distractors.

There is accumulating evidence that where individuals have significant deficits in the 'signal and noise' task, they can detect the direction of motion in the equivalent noise version (Hess et al., 2010; Shepherd, 2006; Tibber et al., 2014) at the same level as the control populations. This suggests that the mechanism used to detect motion coherence between the two tasks differs, and does not require the same process of integration and segmentation required by the signal and noise task. Mansouri and Hess (2006) argue that there is a fundamental difference in tasks that require detecting a coherent signal within visual noise and tasks that require averaging over a mean direction. While both tasks are global in nature, signal and noise tasks require actively pooling the relevant targets that contribute to the signal, while also segmenting out and ignoring the neural response for the noise, while the equivalent noise tasks requires summing the motion direction.

With the signal and noise version, the signal direction is usually defined beforehand. The observer will be advised to attend to a pair of directions (up/down or left/right etc). Since there is an expectation for the signal it is possible that learning may occur more readily for this task as a result of up-weighting for those receptive fields tuned to those directions, based on (i) probabilistic inference at the sensory processing stages (Bejjanki et al., 2011) and (ii) the reentrant (feedback) connections from the global to local levels, which would increase the responses in those channels contributing to the signal (Romei et al., 2016a). In contrast, performance on the equivalent noise tasks requires summing the dots across all the directions. Neither probabilistic inference nor the up-weighting of neural populations would aide in efficiently encoding the mean direction.

In Chapter 4 we trained groups of observers on spatial frequency specific equivalent noise global motion stimuli, that were presented using MOCS, included easy and difficult trials interleaved in addition to external trial-by-trial feedback. There were three spatial frequency conditions (Broad, Low and High spatial frequency) each of the groups was trained on one condition and tested on all three. We tested for transfer between spatial frequencies of global motion, and transfer from the global level to the local level for contrast sensitivity and orientation discrimination. For two of the three conditions performance after training returned to baseline levels, and while this wasn't the purpose of the study it was a particularly interesting finding. It is possible that this can be explained by the up-weighting previously mentioned. This is an interesting area for further investigation, especially in terms of identifying the types of training that may elicit persistent improvements for these tasks. The group that did maintain their improvement after training, was the group trained on low spatial frequency targets. Furthermore this group also showed improved performance for the contrast sensitivity task, at low spatial frequencies. We speculated that this may be as a result of the high correlations in the low spatial frequency channels between the global and local levels.

7.4 Conclusion

Perception and perceptual learning occur as a "whole brain effect" (Maniglia & Seitz, 2018) that includes low level sensory processes (Bejjanki et al., 2011), high-level processes such as attention (Donovan & Carrasco, 2015), confidence (Asher & Hibbard, 2018; Talluri et al., 2015), precision (Jeter et al., 2009), the read-out of sensory inputs (Law & Gold, 2009; Petrov et al., 2005), probabilistic inference (Bejjanki et al., 2011), metacognition (Zylberberg et al., 2014), and internal and external reinforcement (Liu et al., 2012a). We explored a wide range of perceptual tasks including form, motion and depth, and perceptual learning of these tasks. It can be difficult to obtain complete, consistent and reliable data for these measures, particularly in populations such as people who experience migraines. We have demonstrated that by using a statistical model that embraces individual data points that we can include all of our data into our analysis. To that extent, we join a growing number of researchers using mixed effects models to analyse psychophysical data at the sample level.

Using these methods, our studies suggest a crucial role for the reentrant connections from higher to lower areas for perceptual learning and transfer of learning to untrained tasks. Furthermore, from the studies using equivalent noise global motion, there are important questions to ask around the processes of perceiving and improving on the two types of task. In an early computational model Vaina et al. (1995) suggest that improvement in motion coherence detection reflect an increased ability to ignore noise. Down-weighting noise would be consistent with our speculative proposal of relatively up-weighting the signal. We found that self-reported confidence was highly correlated with performance, and consider it a valuable addition to general psychophysical methods. A key aspect to performance and improvement was the level of confidence in the perceptual decision. Finally our large multi-level study suggests that internal reinforcement is as effective as external feedback when difficult trials are interleaved with easy exemplars. Furthermore, in local tasks where there is no added noise, internal feedback was more effective than external feedback. This may suggest that when internal and external confidence conflict at sub-threshold stimulus levels, where uncertainty is high, that the presence of external feedback undermines confidence.

References

Aberg, K. C. & Herzog, M. H. (2009). Interleaving bisection stimuli - randomly or in sequence - does not disrupt perceptual learning, it just makes it more difficult. *Vision Research*, 49(21), 2591–2598.

Adab, Z. H. & Vogels, R. (2011). Practicing coarse orientation discrimination improves orientation signals in macaque cortical area V4. *Current Biology*, 21(19), 1661–1666.

Adelson, E. H. & Movshon, J. A. (1982). Phenomenal coherence of moving visual patterns. *Nature*, 300(5892), 523–525.

Adini, Y., Sagi, D., & Tsodyks, M. (2002). Context-enabled learning in the human visual system. *Nature*, 415(February), 2–5.

Adini, Y., Wilkonsky, A., Haspel, R., Tsodyks, M., & Sagi, D. (2004). Perceptual learning in contrast discrimination: the effect of contrast uncertainty. *Journal of vision*, 4(2004), 993–1005.

Adolph, K. E. & Kretch, K. S. (2015). Learning to Perceive or Perceiving to Learn? *International Encyclopedia of the Social and Behavioral Sciences*, (pp. 127–134).

Aggio, J. F. & Derby, C. D. (2010). Aplysia. In *Encylopedia of Animal Behaviour* (pp. 107–111).

Agresti, A. (2007). *An Introduction to Categorical Data Analysis*, volume 103. Wiley-Interscience.

Agresti, A. (2013). *Categorical Data Analysis*. Hoboken: John Wiley & Sons, 3rd edition.

Ahissar, M. (2007). Dyslexia and the anchoring-deficit hypothesis. *Trends in Cognitive Sciences*, 11(11), 458–465.

Ahissar, M. & Hochstein, S. (1997). Task difficulty and the specificity of perceptual learning. *Letters to Nature*, 387(6631), 401–406.

Ahissar, M. & Hochstein, S. (2004). The reverse hierarchy theory of visual perceptual learning. *Trends in Cognitive Sciences*, 8(10), 457–464.

Ahissar, M., Nahum, M., Nelken, I., & Hochstein, S. (2009). Reverse hierarchies and sensory learning. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1515), 285–299.

Ahmed, N. S., Gottlob, I., Proudlock, F. A., & Hutchinson, C. V. (2018). Restricted Spatial Windows of Visibility in Myalgic Encephalomyelitis (ME). *Vision*, 2(1), 2.

Allenmark, F. & Read, J. C. A. (2011). Spatial stereoresolution for depth corrugations may be set in primary visual cortex. *BMC Neuroscience*, 12(1), P263.

Allison, P. D. (2010). Missing data. In P. V. Marsden & D. Write, James (Eds.), *Handbook of Survey Research*, volume 136 chapter Chapter20, (pp. 631–657). Emerald Group Publishing, second edition.

Amano, K., Edwards, M., Badcock, D. R., & Nishida, S. (2009). Spatial-frequency tuning in the pooling of one- and two-dimensional motion signals. *Vision Research*, 49(23), 2862–2869.

Amitay, S., Irwin, A., & Moore, D. R. (2006). Discrimination learning induced by training with identical stimuli. *Nature neuroscience*, 9(11), 1446–1448.

An, X., Gong, H., Qian, L., Wang, X., Pan, Y., Zhang, X., Yang, Y., & Wang, W. (2012). Distinct Functional Organizations for Processing Different Motion Signals in V1, V2, and V4 of Macaque. *Journal of Neuroscience*, 32(39), 13363–13379.

Anstis, S. M. & Rogers, B. J. (1975). Illusory reversal of visual depth and movement during changes of contrast. *Vision research*, 15(8), 957–IN6.

Antal, A., Temme, J., Nitsche, M. A., Varga, E. T., Lang, N., & Paulus, W. (2005). Altered motion perception in migraineurs: evidence for interictal cortical hyperexcitability. *Cephalalgia*, 25(10), 788–794.

Anzai, A., Chowdhury, S. A., & DeAngelis, G. C. (2011). Coding of stereoscopic depth information in visual areas V3 and V3A. *Journal of Neuroscience*, 31(28), 10270–10282.

Aoki, S. C., Shiozaki, H. M., & Fujita, I. (2017). A relative frame of reference underlies reversed depth perception in anticorrelated random-dot stereograms. *Journal of Vision*, 17(12), 17.

Asher, J. M. & Hibbard, P. B. (2018). First- and second-order contributions to depth perception in anti-correlated random dot stereograms. *Scientific Reports*, 8(14120), 1–20.

Asher, J. M., O'Hare, L., Romei, V., & Hibbard, P. B. (2018a). Typical Lateral Interactions, but Increased Contrast Sensitivity, in Migraine-With-Aura. *Vision*, 2(1), 7.

Asher, J. M., Romei, V., & Hibbard, P. B. (2018b). Spatial frequency tuning of perceptual learning and transfer in global motion. *bioRxiv*.

Assee, A. & Qian, N. (2007). Solving da Vinci stereopsis with depth-edge-selective V2 cells. *Vision research*, 47(20), 2585–2602.

Astle, A. T., Li, R. W., Webb, B. S., Levi, D. M., & McGraw, P. V. (2013). A Weber-like law for perceptual learning. *Scientific Reports*, 3, 1–8.

Astle, A. T., McGraw, P. V., & Webb, B. S. (2011a). Can human amblyopia be treated in adulthood? *Strabismus*, 19(3), 99–109.

Astle, A. T., Webb, B. S., & McGraw, P. V. (2011b). The pattern of learned visual improvements in adult amblyopia. *Investigative Ophthalmology and Visual Science*, 52(10), 7195–7204.

Aurora, S. K. & Wilkinson, F. (2007). The brain is hyperexcitable in migraine. *Cephalalgia*, 27(12), 1442–1453.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59(4), 390–412.

Baker, D. H. & Vilidaite, G. (2014). Broadband noise masks suppress neural responses to narrowband stimuli. *Frontiers in Psychology*, 5(JUL), 1–9.

Baker, D. H., Wallis, S. A., Georgeson, M. A., & Meese, T. S. (2012). The effect of interocular phase difference on perceived contrast. *PLoS One*, 7(4), e34696.

Baldassarre, A., Lewis, C. M., Committeri, G., Snyder, A. Z., Romani, G. L., & Corbetta, M. (2012). Individual variability in functional connectivity predicts performance of a perceptual task. *Proceedings of the National Academy of Sciences*, 109(9), 3516–3521.

Baldassarree, A., Lewis, C. M., Committeri, G., Snyder, A. Z., Romani, G. L., & Corbetta, M. (2016). Correction for Baldassarre et al., Individual variability in functional connectivity predicts performance of a perceptual task. *Proceedings of the National Academy of Sciences*, 113(43), E6723–E6723.

Ball, K. & Sekuler, R. (1982). A specific and enduring improvement in visual motion discrimination. *Science*, 218(4573), 697–698.

Ball, K. & Sekuler, R. (1987). Direction-specific improvement in motion discrimination. *Vision Research*, 27(6), 953–965.

Banks, M. S., Aslin, R. N., Letson, R. D., Url, S., & Rose, H. N. (1975). Sensitive Period for the Development of Human Binocular Vision. *Science*, 190(4215), 675–677.

Banks, M. S., Gepshtein, S., & Landy, M. S. (2004). Why is spatial stereoresolution so low? *Journal of Neuroscience*, 24(9), 2077–2089.

Barlow, H. (1957). Increment Thresholds at low intensities considered as signal / noise disrciminations. *Journal of Physiology*, 136(1760), 469–488.

Barlow, H. (1990). Conditions for versatile learning, Helmholtz's unconscious inference, and the task of perception. *Vision Research*, 30(11), 1561–1571.

Barlow, H. & Foldiak, P. (1989). The computing neuron. *Adaptation and decorrelation in the cortex*, (pp. 54–72).

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3), 255–278.

Baseler, H. A., Gouws, A., Haak, K. V., Racey, C., Crossland, M. D., Tufail, A., Rubin, G. S., Cornelissen, F. W., & Morland, A. B. (2011). Large-scale remapping of visual cortex is absent in adult humans with macular degeneration. *Nature neuroscience*, 14(5), 649–655.

Battista, J., Badcock, D. R., & McKendrick, A. M. (2011). Migraine increases centre-surround suppression for drifting visual stimuli. *PLoS ONE*, 6(4).

Bejjanki, V. R., Beck, J. M., Lu, Z.-L., & Pouget, A. (2011). Perceptual learning as improved probabilistic inference in early sensory areas. *Nature Neuroscience*, 14(5), 642–50.

Benedek, K., Tajti, J., Janáky, M., Vécsei, L., & Benedek, G. (2002). Spatial contrast sensitivity of migraine patients without aura. *Cephalalgia*, 22(2), 142–145.

Bex, P. J. & Dakin, S. C. (2002). Comparison of the spatial-frequency selectivity of local and global motion detectors. *Journal of the Optical Society of America*, 19(4), 670–7.

Biederman, I. & Shiffrar, M. M. (1987). Sexing Day-Old Chicks - a Case-Study and Expert Systems- Analysis of a Difficult Perceptual-Learning Task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13(4), 640–645.

Blake, R. & Wilson, H. (2011). Binocular vision. *Vision Research*, 51(7), 754–770.

Bolker, B. M., Brooks, M. E., Clark, C. J., Geange, S. W., Poulsen, J. R., Stevens, M. H. H., & White, J. S. S. (2009). Generalized linear mixed models: a practical guide for ecology and evolution. *Trends in Ecology and Evolution*, 24(3), 127–135.

Born, R. T. & Bradley, D. C. (2005). Structure and Function of Visual Area Mt. *Annual Review of Neuroscience*, 28(1), 157–189.

Bosten, J. M., Bargary, G., Goodbourn, P. T., Hogg, R. E., Lawrance-Owen, A. J., & Mollon, J. (2014). Individual differences provide psychophysical evidence for separate on- and off-pathways deriving from short-wave cones. *Journal of the Optical Society of America. A, Optics, image science, and vision*, 31(4), A47–A54.

Braddick, O. J. (1993). Segmentation versus integration in visual motion processing. *Trends in Neurosciences*, 16(7), 263–268.

Braddick, O. J. (1997). Local and global representations of velocity: Transparency, opponency, and global direction perception. *Perception*, 26(8), 995–1010.

Braddick, O. J., Atkinson, J., & Wattam-Bell, J. (2003). Normal and anomalous development of visual motion processing: Motion coherence and 'dorsal-stream vulnerability'. *Neuropsychologia*, 41(13), 1769–1784.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial vision*, 10, 433–436.

Braunitzer, G., Rokszin, A., Kóbor, J., & Benedek, G. (2010). Is the development of visual contrast sensitivity impaired in children with migraine? An exploratory study. *Cephalalgia : an international journal of headache*, 30(8), 991–5.

Bredfeldt, C. E. & Cumming, B. G. (2006). A simple account of cyclopean edge responses in macaque V2. *Journal of Neuroscience*, 26(29), 7581–7596.

Bredfeldt, C. E., Read, J. C. A., & Cumming, B. G. (2009). A quantitative explanation of responses to disparity-defined edges in macaque V2. *Journal of neurophysiology*, 101(2), 701–713.

Breitmeyer, B. G. & Ogmen, H. (2000). Recent models and findings in visual backward masking: A comparison, review, and update. *Perception & psychophysics*, 62(8), 1572–1595.

Bridge, H., Stagg, C. J., Near, J., Lau, C.-i., Zisner, A., & Cader, M. Z. (2015). Altered neurochemical coupling in the occipital cortex in migraine with visual aura. *Cephalalgia*, 35(11), 1025–1030.

Britten, K. H., Shadlen, M. N., Newsome, W. T., & Movshon, J. A. (1992). The analysis of visual motion: a comparison of neuronal and psychophysical performance. *The Journal of Neuroscience*, 12(12), 4745–4765.

Britten, K. H., Shadlen, M. N., Newsome, W. T., & Movshon, J. A. (1993). Responses of neurons in macaque MT to stochiastic motion signals.

Bruce, V., Green, P. R., & Georgeson, M. A. (2003). Vision for Awareness : Approaches to the Psychology of Visual Perception. In 4th (Ed.), *Visual perception: Physiology, psychology, & ecology* chapter 4, (pp. 77–83). Psychology Press.

Burch, R. C., Loder, S., Loder, E., & Smitherman, T. A. (2015). The prevalence and burden of migraine and severe headache in the United States: updated statistics from government health surveillance studies. *Headache: The Journal of Head and Face Pain*, 55(1), 21–34.

Burr, D. & Thompson, P. (2011). Motion psychophysics: 1985-2010. *Vision Research*, 51(13), 1431–1456.

Byrne, J. H. & Hawkins, R. D. (2015). Nonassociative learning in invertebrates. *Cold Spring Harbor Perspectives in Biology*, 7(5), 1–16.

Caelli, T. & Porter, D. (1980). On difficulties in localizing ambulance sirens. *Human Factors*, 22(6), 719–724.

Callaway, E. M. (2004). Feedforward, feedback and inhibitory connections in primate visual cortex. *Neural Networks*, 17(5-6), 625–632.

Callaway, E. M. & Nassi, J. J. (2009). Parallel processing strategies of the primate visual system. *Nature Reviews Neuroscience*, 10(5), 360–372.

Camilleri, R., Pavan, A., Ghin, F., Battaglini, L., & Campana, G. (2014a). Improvement of uncorrected visual acuity (UCVA) and contrast sensitivity (UCCS) with perceptual learning and transcranial random noise stimulation (tRNS) in individuals with mild myopia. *Frontiers in Psychology*, 5(OCT), 1–6.

Camilleri, R., Pavan, A., Ghin, F., & Campana, G. (2014b). Improving myopia via perceptual learning: is training with lateral masking the only (or the most) efficacious technique? *Attention Perception & Psychophysics*, 76, 2485–2494.

Campana, G. & Maniglia, M. (2015). Editorial: Improving visual deficits with perceptual learning. *Frontiers in Psychology*, 6(APR), 1–3.

Casals, M., Girabent-Farrés, M., & Carrasco, J. L. (2014). Methodological quality and reporting of generalized linear mixed models in clinical medicine (2000-2012): A systematic review. *PLoS ONE*, 9(11), 1–10.

Catchpole, K. & Mckeown, D. (2007). A framework for the design of ambulance sirens. *Ergonomics*, 50(8), 1287–1301.

Chiappini, E., Silvanto, J., Hibbard, P. B., Avenanti, A., & Romei, V. (2018). Strengthening functionally specific neural pathways with transcranial brain stimulation. *Current Biology*, 28(13), R735–R736.

Chu, M. W., Li, W. L., & Komiyama, T. (2016). Balancing the Robustness and Efficiency of Odor Representations during Learning. *Neuron*, 92(1), 174–186.

Chubb, C., Sperling, G., & Solomon, J. A. (1989). Texture interactions determine perceived contrast. *Proceedings of the National Academy of Sciences*, 86(23), 9631–9635.

Coates, D. R. & Chung, S. T. (2014). Changes across the psychometric function following perceptual learning of an RSVP reading task. *Frontiers in Psychology*, 5(DEC), 1–9.

Cogan, A. I., Lomakin, A. J., & Rossi, A. F. (1993). Depth in anticorrelated stereograms: effects of spatial density and interocular delay. *Vision research*, 33(14), 1959–1975.

Cong, L.-J., Wang, R.-J., Yu, C., & Zhang, J.-Y. (2016). Perceptual learning of basic visual features remains task specific with Training-Plus-Exposure (TPE) training. *Journal of Vision*, 16(3:13), 1–9.

Cong, L.-J. & Zhang, J.-Y. (2014). Perceptual learning of contrast discrimination under roving: The role of semantic sequence in stimulus tagging. *Journal of Vision*, 14(13), 1–8.

Cook, M. & Gillam, B. (2004). Depth of Monocular Elements in a Binocular Scene: The Conditions for da Vinci Stereopsis. *Journal of Experimental Psychology: Human Perception and Performance*, 30(1), 92–103.

Coppola, G., Parisi, V., Di Lorenzo, C., Serrao, M., Magis, D., Schoenen, J., & Pierelli, F. (2013). Lateral inhibition in visual cortex of migraine patients between attacks. *The journal of headache and pain*, 14(1), 20.

Coppola, G., Pierelli, F., & Schoenen, J. (2007). Is the cerebral cortex hyperexcitable or hyperresponsive in migraine? *Cephalalgia*, 27(12), 1427–1439.

Cormack, L. K., Stevenson, S. B., & Schor, C. M. (1991). Interocular correlation, luminance contrast and cyclopean processing. *Vision research*, 31(12), 2195–2207.

Cottereau, B. R., McKee, S. P., & Norcia, A. M. (2012). Bridging the gap: Global disparity processing in the human visual cortex. *Journal of Neurophysiology*, 107(9), 2421–2429.

Cowey, A., Campana, G., Walsh, V., & Vaina, L. M. (2006). The role of human extra-striate visual areas V5/MT and V2/V3 in the perception of the direction of global motion: A transcranial magnetic stimulation study. *Experimental Brain Research*, 171(4), 558–562.

Croner, L. J. & Kaplan, E. (1995). Receptive fields of P and M ganglion cells across the primate retina. *Vision Research*, 35(1), 7–24.

Cumming, B. G. & Parker, A. J. (1997). Responses of primary visual cortical neurons to binocular disparity without depth perception. *Nature*, 389(6648), 280–283.

Cumming, B. G. & Parker, A. J. (1999). Binocular neurons in V1 of awake monkeys are selective for absolute, not relative, disparity. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 19(13), 5602–18.

Cumming, B. G., Shapiro, S. E., & Parker, A. J. (1998). Disparity detection in anticorrelated stereograms. *Perception*, 27(11), 1367–1377.

Dakin, S. C. & Bex, P. J. (2001). Local and global visual grouping: Tuning for spatial frequency and contrast. *Journal of Vision*, 1(2001), 99–111.

Dakin, S. C., Mareschal, I., & Bex, P. J. (2005a). An oblique effect for local motion: psychophysics and natural movie statistics. *Journal of vision*, 5(10), 878–87.

Dakin, S. C., Mareschal, I., & Bex, P. J. (2005b). Local and global limitations on direction integration assessed using equivalent noise analysis. *Vision Research*, 45(24), 3027–3049.

Das, A. & Huxlin, K. R. (2010). New Approaches to Visual Rehabilitation for Cortical Blindness: Outcomes and Putative Mechanisms. *The Neuroscientist*, 16(4), 374–387.

Das, a., Tadin, D., & Huxlin, K. R. (2014). Beyond Blindsight: Properties of Visual Relearning in Cortically Blind Fields. *Journal of Neuroscience*, 34(35), 11652–11664.

Daw, N. W. (2009). The foundations of development and deprivation in the visual system. *Journal of Physiology*, 587(12), 2769–2773.

DeValois, R. L., Albrecht, D. G., & Thonrell, L. G. (1982). Spatial Frequency Selectivity of Cells in Macaque Visual Cortex. *Vision research*, 22(5), 545–559.

Deveau, J., Ozer, D. J., & Seitz, A. R. (2014). Improved vision and on-field performance in baseball through perceptual learning. *Current Biology*, 24(4), R146–R147.

Ding, J. & Levi, D. M. (2011). Recovery of stereopsis through perceptual learning in human adults with abnormal binocular vision. *Proceedings of the National Academy of Sciences*, 108(37), E733–E741.

Dinse, H. R., Ragert, P., Pleger, B., Schwenkreis, P., & Tegenthoff, M. (2003). Pharmacological Modulation of Perceptual Learning and Associated Cortical Reorganization TL - 301. *Science*, 301 VN -(5629), 91–94.

Ditchfield, J. A., McKendrick, A. M., & Badcock, D. R. (2006). Processing of global form and motion in migraineurs. *Vision Research*, 46(1-2), 141–148.

Dobres, J. & Watanabe, T. (2012). Response feedback triggers long-term consolidation of perceptual learning independently of performance gains. *Journal of Vision*, 12(8), 9–9.

Doi, T. & Fujita, I. (2014). Cross-matching: a modified cross-correlation underlying threshold energy model and match-based depth perception. *Frontiers in computational neuroscience*, 8(October), 127.

Doi, T., Tanabe, S., & Fujita, I. (2011). Matching and correlation computations in stereoscopic depth perception. *Journal of Vision*, 11(3), 1.

Donovan, I. & Carrasco, M. (2015). Exogenous attention facilitates location transfer of perceptual learning. *Journal of vision*, 15(10), 1–16.

Dosher, B. A., Jeter, P. E., Liu, J., & Lu, Z.-L. (2013). An integrated reweighting theory of perceptual learning. *Proceedings of the National Academy of Sciences of the United States of America*, 110(33), 13678–13683.

Dosher, B. A. & Lu, Z. L. (1998). Perceptual learning reflects external noise filtering and internal noise reduction through channel reweighting. *Proceedings of the National Academy of Sciences of the United States of America*, 95(23), 13988–93.

Dosher, B. A. & Lu, Z. L. (1999). Mechanisms of perceptual learning. *Vision Research*, 39, 3197–3221.

Dosher, B. A. & Lu, Z. L. (2009). Hebbian Reweighting on Stable Representations in Perceptual Learning. *Learning & Perception*, 1(1), 37–58.

Dosher, B. A. & Lu, Z.-L. (2017). Visual Perceptual Learning and Models. *Annual Review of Vision Science*, 3(1), 1–24.

Doyon, J. & Benali, H. (2005). Reorganization and plasticity in the adult brain during learning of motor skills. *Current Opinion in Neurobiology*, 15(2), 161–167.

Durgin, F. H., Srimant, T. P., & Levi, D. M. (1995). On the filling in of the visual blind spot: some rules of thumb. *Perception*, 24(7), 827–840.

Edden, R. A. E., Muthukumaraswamy, S. D., Freeman, T. C. A., & Singh, K. D. (2009). Orientation discrimination performance is predicted by GABA concentration and gamma oscillation frequency in human primary visual cortex. *Journal of Neuroscience*, 29(50), 15721–15726.

Edwards, M., Cassanello, C. R., Badcock, D. R., & Nishida, S. (2013). Effect of form cues on 1D and 2D motion pooling. *Vision Research*, 76, 94–104.

Edwards, M. & Metcalf, O. (2010). Independence in the processing of first- and second-order motion signals at the local-motion-pooling level. *Vision Research*, 50(3), 261–270.

Edwards, M. & Nishida, S. (2004). Contrast-reversing global-motion stimuli reveal local interactions between first-and second-order motion signals. *Vision research*, 44(16), 1941–1950.

Edwards, M., Pope, D. R., & Schor, C. M. (1999). Orientation tuning of the transient-stereopsis system. *Vision Research*, 39(16), 2717–2727.

Edwards, M., Pope, D. R., & Schor, C. M. (2000). First- and second-order processing in transient stereopsis. *Vision Research*, 40(19), 2645–2651.

Erkelens, C. J. & Collewijn, H. (1985a). Eye movements and stereopsis during dichoptic viewing of moving random-dot stereograms. *Vision research*, 25(11), 1689–1700.

Erkelens, C. J. & Collewijn, H. (1985b). Motion perception during dichoptic viewing of moving random-dot stereograms. *Vision research*, 25(4), 583–588.

Fahle, M. (1997). Specificity of learning curvature, orientation, and vernier discriminations. *Vision Research*, 37(14), 1885–1895.

Fahle, M. (2002). Introduction. In M. Fahle & T. Poggio (Eds.), *Perceptual Learning* (pp. ix–xx). MIT Press, 1 edition.

Fahle, M. (2005). Perceptual learning: Specificity versus generalization. *Current Opinion in Neurobiology*, 15(2), 154–160.

Fahle, M. & Edelman, S. (1993). Long-term learning in vernier acuity: Effects of stimulus orientation, range and of feedback. *Vision Research*, 33(3), 397–412.

Fahle, M., Edelman, S., & Poggio, T. (1995). Fast perceptual learning in hyperacuity. *Vision Research*, 35(21), 3003–3013.

Fahle, M. & Henke-Fahle, S. (1996). Interobserver variance in perceptual performance and learning. *Investigative Ophthalmology and Visual Science*, 37(5), 869–877.

Felleman, D. J. & Van Essen, D. C. (1987). Receptive field properties of neurons in area V3 of macaque monkey extrastriate cortex. *Journal of neurophysiology*, 57(4), 889–920.

Felleman, D. J. & Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex*, 1(1), 1–47.

Fesharaki, H., Rezaei, L., Farrahi, F., Banihashem, T., & Jahanbkhshi, A. (2012). Normal interpupillary distance values in an Iranian population. *Journal of Ophthalmic and Vision Research*, 7(3), 231–234.

Field, D. J., Hayes, A., & Hess, R. F. (1993). Contour integration by the human visual system: evidence for a local “association field”. *Vision research*, 33(2), 173–193.

Fine, I. & Jacobs, R. A. (2002). Comparing perceptual learning tasks: a review. *Journal of Vision*, 2(2), 190–203.

Fiorentini, A. & Berardi, N. (1980). Perceptual learning specific for orientation and spatial frequency. *Nature*, 287, 43–44.

Fleet, D. J., Wagner, H., & Heeger, D. J. (1996). Neural Encoding of Binocular Disparity: Energy Models, Position Shifts and Phase Shifts. *Vision Research*, 36(12), 1839–1857.

Furlan, M. & Smith, A. T. (2016). Global Motion Processing in Human Visual Cortical Areas V2 and V3. *Journal of Neuroscience*, 36(27), 7314–7324.

Furmanski, C. S., Schluppeck, D., & Engel, S. A. (2014). Learning strengthens the response of primary visual cortex to simple patterns. *Current Biology*, 14(7), 573–578.

Garcia, A., Kuai, S. G., & Kourtzi, Z. (2013). Differences in the time course of learning for hard compared to easy training. *Frontiers in Psychology*, 4(MAR), 1–8.

García-Palacios, A., Hoffman, H., Carlin, A., Furness III, T. A., & Botella, C. (2002). Virtual reality in the treatment of spider phobia: a controlled study. *Behaviour Research and Therapy*, 40, 983–993.

García-Pérez, M. A. & Alcalá-Quintana, R. (2005). Sampling plans for fitting the psychometric function. *Spanish Journal of Psychology*, 8(2), 256–289.

Gawne, T. J. (2002). Responses of Primate Visual Cortical Neurons to Stimuli Presented by Flash, Saccade, Blink, and External Darkening. *Journal of Neurophysiology*, 88(5), 2178–2186.

Gescheider, G. A. (1997). *Psychophysics: The Fundamentals*. Psychology Press, 3 edition.

Gibson, E. J. (1963). Perceptual Learning. *Annual review of psychology*, 14(October), 29–53.

Gibson, J. J. (2014). *The Ecological Approach to Visu*. New York, NY: Psychology Press, 1st edition.

Gibson, J. J. & Gibson, E. J. (1955). Perceptual learning: Differentiation or enrichment? *Psychological Review*, 62(1), 32–41.

Gilbert, C. D. (1998). Adult Cortical Dynamics. *Physiological Reviews*, 78(2), 467–485.

Gilbert, C. D. & Li, W. (2013). Top-down influences on visual processing. *Nature Reviews Neuroscience*, 14(5), 350–363.

Gilbert, C. D. & Sigman, M. (2007). Brain states: top-down influences in sensory processing. *Neuron*, 54(5), 677–696.

Gilbert, C. D., Sigman, M., & Crist, R. E. (2001). The Neural Basis of Perceptual Learning Review. *Neuron*, 31, 681–697.

Gilbert, C. D. & Wiesel, T. N. (1992). Receptive field dynamics in adult primary visual cortex. *Nature*, 356(March), 150–152.

Gillam, B. & Nakayama, K. (1999). Quantitative depth for a phantom surface can be based on cyclopean occlusion cues alone. *Vision Research*, 39(1), 109–112.

Glass, L. (1969). Moire effect from random dots. *Nature*, 223(5206), 578.

Gold, J. I., Law, C.-T., Connolly, P., & Bennur, S. (2010). Relationships Between the Threshold and Slope of Psychometric and Neurometric Functions During Perceptual Learning: Implications for Neuronal Pooling. *Journal of Neurophysiology*, 103(1), 140–154.

Gold, J. I., Sekuler, A. B., & Bennett, P. J. (2004). Characterizing perceptual learning with external noise. *Cognitive Science*, 28(2), 167–207.

Gold, J. I. & Watanabe, T. (2010). Perceptual learning. *Current Biology*, 20(2), 1–4.

Goodale, M. A. & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in neurosciences*, 15(1), 20–5.

Grill-Spector, K., Kourtzi, Z., & Kanwisher, N. (2001). The lateral occipital complex and its role in object recognition. *Vision Research*, 41, 1409–1422.

Gu, Y., Liu, S., Fetsch, C. R., Yang, Y., Fok, S., Sunkara, A., DeAngelis, G. C., & Angelaki, D. E. (2011). Perceptual learning reduces interneuronal correlations in macaque visual cortex. *Neuron*, 71(4), 750–761.

Gur, M. & Snodderly, D. M. (2007). Direction selectivity in V1 of alert monkeys: Evidence for parallel pathways for motion processing. *Journal of Physiology*, 585(2), 383–400.

Haak, K. V., Morland, A. B., & Engel, S. A. (2015). Plasticity, and Its Limits, in Adult Human Primary Visual Cortex. *Multisensory research*, 28(3-4), 297–307.

Hadid, V. & Lepore, F. (2017). From Cortical Blindness to Conscious Visual Perception: Theories on Neuronal Networks and Visual Training Strategies. *Frontiers in Systems Neuroscience*, 11(August), 1–10.

Hall, G. (1991). Associative theory and the phenomena of perceptual learning. In *Perceptual and Associative Learning* (pp. 1–27). Oxford University Press, 1 edition.

Hall, G. (2003). Learned changes in the sensitivity of stimulus representations: Associative and nonassociative mechanisms. *Quarterly Journal of Experimental Psychology*, 56 B(1), 43–55.

Halpern, S. D., Andrews, T. J., & Purves, D. (1999). Interindividual variation in human visual performance. *Journal of cognitive neuroscience*, 11(5), 521–34.

Harris, H., Gliksberg, M., & Sagi, D. (2012). Generalized perceptual learning in the absence of sensory adaptation. *Current Biology*, 22(19), 1813–1817.

Harris, J. M., Nefs, H. T., & Grafton, C. E. (2008). Binocular vision and motion-in-depth. *Spatial Vision*, 21(6), 531–547.

Harrison, R. V., Gordon, K. A., & Mount, R. J. (2005). Is there a critical period for cochlear implantation in congenitally deaf children? Analyses of hearing and speech perception performance after implantation. *Developmental Psychobiology*, 46(3), 252–261.

Harrison, X. A., Donaldson, L., Correa-Cano, M. E., Evans, J., Fisher, D. N., Goodwin, C. E., Robinson, B. S., Hodgson, D. J., & Inger, R. (2018). A brief introduction to mixed effects modelling and multi-model inference in ecology. *PeerJ*, 6, e4794.

Hawken, M. J., Parker, A. J., & Lund, J. S. (1988). Laminar organization and contrast sensitivity of direction-selective cells in the striate cortex of the Old World monkey. *Journal of Neuroscience*, 8(10), 3541–3548.

Hawkins, R. D. & Byrne, J. H. (2015a). Associative learning in invertebrates. *Cold Spring Harbor Perspectives in Biology*, 7(5), 1–18.

Hawkins, R. D. & Byrne, J. H. (2015b). Associative learning in invertebrates. *Cold Spring Harbor Perspectives in Biology*, 7(5), 1–18.

Headache Classification Committee of the International Headache Society (2013). Classification and diagnostic criteria for headache disorders, cranial neuralgias and facial pain. *Cephalalgia*, 33(9), 629–808.

Hedges, J. H., Gartshteyn, Y., Kohn, A., Rust, N. C., Shadlen, M. N., Newsome, W. T., & Movshon, J. A. (2011). Dissociation of neuronal and psychophysical responses to local and global motion. *Current Biology*, 21(23), 2023–2028.

Heinen, S. J. & Skavenski, A. (1991). Recovery of visual responses in foveal V1 neurons following bilateral foveal lesions in adult monkey. *Experimental Brain Research*, 83(September), 670–674.

Henriksen, S., Cumming, B. G., & Read, J. C. A. (2016). A Single Mechanism Can Account for Human Perception of Depth in Mixed Correlation Random Dot Stereograms. *PLoS Computational Biology*, 12(5), 1–21.

Henriksson, L., Nurminen, L., Hyvarinen, A., Vanni, S., Hyvarinen, A., Vanni, S., Hyvarinen, A., & Vanni, S. (2008). Spatial frequency tuning in human retinotopic visual areas. *Journal of Vision*, 8(2008), 1–13.

Hensch, T. K. (2004). Critical Period Regulation. *Annual Review of Neuroscience*, 27(1), 549–579.

Hensch, T. K. (2005). Critical period plasticity in local cortical circuits. *Nature Reviews Neuroscience*, 6(11), 877–888.

Heron, S. & Lages, M. (2012). Screening and sampling in studies of binocular vision. *Vision Research*, 62, 228–234.

Herrmann, K., Montaser-Kouhsari, L., Carrasco, M., & Heeger, D. J. (2010). When size matters: Attention affects performance by contrast or response gain. *Nature Neuroscience*, 13(12), 1554–1561.

Herzog, M. H., Aberg, K. C., Frémaux, N., Gerstner, W., & Sprekeler, H. (2012). Perceptual learning, roving and the unsupervised bias. *Vision Research*, 61, 95–99.

Herzog, M. H., Ewald, K. R. F., Hermens, F., & Fahle, M. (2006). Reverse feedback induces position and orientation specific changes. *Vision Research*, 46(22), 3761–3770.

Herzog, M. H. & Fahle, M. (1997). The Role of Feedback in Learning a Vernier Discrimination Task. *Vision Research*, 37(15), 2133–2141.

Herzog, M. H. & Fahle, M. (1998). Modeling perceptual learning: difficulties and how they can be overcome. *Biological cybernetics*, 78(2), 107–117.

Herzog, M. H. & Fahle, M. (1999). A recurrent model for perceptual learning. *Journal of Optical Technology*, 66(9), 836.

Hess, R. F., Hayes, A., & Field, D. J. (2003). Contour integration and cortical processing. *Journal of Physiology Paris*, 97(2-3), 105–119.

Hess, R. F., Mansouri, B., & Thompson, B. (2010). A binocular approach to treating amblyopia: Antisuppression therapy. *Optometry and Vision Science*, 87(9), 697–704.

Hess, R. F. & Wilcox, L. M. (1994). Linear and non-linear filtering in stereopsis. *Vision research*, 34(18), 2431–2438.

Hibbard, P. B., Bradshaw, M. F., Langley, K., & Rogers, B. J. (2002). The stereoscopic anisotropy: individual differences and underlying mechanisms. *J Exp Psychol Hum Percept Perform*, 28(2), 469–476.

Hibbard, P. B., Goutcher, R., & Hunter, D. (2016). Encoding and estimation of first- and second-order binocular disparity in natural images. *Vision Research*, 120, 108–120.

Hibbard, P. B., Scott-Brown, K., Haigh, E. C., & Adrain, M. (2014). Depth perception not found in human observers for static or dynamic anti-correlated random dot stereograms. *PLoS ONE*, 9(1), e84087.

Hochstein, S. & Ahissar, M. (2002). Hierarchies and reverse hierarchies in the Visual System. *Neuron*, 36(3), 791–804.

Howard, C. Q., Aaron, J. M., & Elefterios, P. P. (2011). Acoustic characteristics for effective ambulance sirens. *Acoustics Australia*, 39(2), 43–53.

Hu, J., Ma, H., Zhu, S., Li, P., Xu, H., Fang, Y., Chen, M., Han, C., Fang, C., Cai, X., Yan, K., & Lu, H. D. (2018). Visual Motion Processing in Macaque V2. *Cell Reports*, 25(1), 157–167.

Huang, X., Lu, H., Tjan, B. S., Zhou, Y., & Liu, Z. (2007). Motion perceptual learning: when only task-relevant information is learned. *Journal of vision*, 7(10), 1–10.

Hubel, D. H. & Wiesel, T. N. (1959). Receptive fields of single neurones in the cat's striate cortex. *The Journal of physiology*, 148(3), 574–591.

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

Hubel, D. H. & Wiesel, T. N. (1965). Receptive fields and functional acrhitecture in two nonstriate visual areas (18 and 19) of the cat. *Journal of neurophysiology*, 28(2), 229–289.

Hubel, D. H. & Wiesel, T. N. (1968). Receptive Fields and Functional Architecture. *October*, 587(Pt 12), 215–243.

Hubel, D. H. & Wiesel, T. N. (1970). The period of susceptibility to the physiological effects of unilateral eye closure in kittens. *Journal of Physiology*, 206, 419–436.

Hubel, D. H. & Wiesel, T. N. (1972). Laminar and columnar distribution of geniculo-cortical fibers in the macaque monkey. *Journal of Comparative Neurology*, 146(4), 421–450.

Hubel, D. H. & Wiesel, T. N. (1977). Functional architecture of macaque monkey visual cortex. *Society*, 198(1130), 1–59.

Hubel, D. H., Wiesel, T. N., & Stryker, M. P. (1978). Anatomical demonstration of orietation columns in macaque monkey. *J Comp Neurol*, 177(3)(PY - 1978), 361–380.

Huber, C. (2014). How to simulate multilevel/longitudinal data.

Hung, S.-C. & Seitz, A. R. (2014). Prolonged Training at Threshold Promotes Robust Retinotopic Specificity in Perceptual Learning. *Journal of Neuroscience*, 34(25), 8423–8431.

Hupé, J.-M., James, A. C., Girard, P., Lomber, S. G., Payne, B. R., & Bullier, J. (2001). Feedback connections act on the early part of the responses in monkey visual cortex. *Journal of Neurophysiology*, 85(1), 134–145.

Huxlin, K. R. (2008). Perceptual plasticity in damaged adult visual systems. *Vision Research*, 48(20), 2154–2166.

Huxlin, K. R., Martin, T., Kelly, K., Riley, M., Friedman, D. I., Burgin, W. S., & Hayhoe, M. (2009). Perceptual relearning of complex visual motion after V1 damage in humans. *The Journal of Neuroscience*, 29(13), 3981–3991.

Janssen, P., Vogels, R., Liu, Y., & Orban, G. A. G. G. A. (2003). At least at the level of inferior temporal cortex, the stereo correspondence problem is solved. *Neuron*, 37(4), 693–701.

Jehee, J. F., Ling, S., Swisher, J. D., van Bergen, R. S., & Tong, F. (2012). Perceptual Learning Selectively Refines Orientation Representations in Early Visual Cortex. *Journal of Neuroscience*, 32(47), 16747–16753.

Jehee, J. F., Roelfsema, P. R., Deco, G., Murre, J. M. J., & Lamme, V. A. F. (2007). Interactions between higher and lower visual areas improve shape selectivity of higher level neurons-Explaining crowding phenomena. *Brain Research*, 1157(1), 167–176.

Jeter, P. E., Dosher, B. A., & Lu, Z.-L. (2009). Task precision at transfer determines specificity of perceptual learning. *Journal of Vision*, 9(3), 1–13.

Julesz, B. (1971). *Foundations of cyclopean perception.* U. Chicago Press.

Kaernbach, C. (2001). Slope bias of psychometric functions derived from adaptive data. *Perception and Psychophysics*, 63(8), 1389–1398.

Kamihirata, H., Oga, T., Aoki, S., & Fujita, I. (2015). A gap between adjacent surfaces deteriorates depth perception based on binocular correlation computation. *Journal of Physiological Sciences*, 17((Suppl. 1)), s155.

Kandel, E. R. (2004). The molecular biology of memory storage: A dialog between genes and synapses. *Bioscience Reports*, 24(4-5), 477–522.

Kandel, E. R. & Spencer, W. A. (1968). Cellular neurophysiological approaches in the study of learning. *Physiological reviews*, 48(1), 65–134.

Kaplan, E. & Shapley, R. M. (1982). X and Y cells in the lateral geniculate nucleus of macaque monkeys. *The Journal of Physiology*, 330(1), 125–143.

Karni, A. & Sagi, D. (1991). Where practice makes perfect in texture discrimination: evidence for primary visual cortex plasticity. *Proceedings of the National Academy of Sciences of the United States of America*, 88(11), 4966–4970.

Karni, A. & Sagi, D. (1993). The time course of learning a visual skill. *Nature*, 365(6443), 250–252.

Kasten, E. & Sabel, B. A. (1995). Visual field enlargement after computer training in brain-damaged patients with homonymous deficits: An open pilot trial. *Restorative Neurology and Neuroscience*, 8(3), 113–127.

Kattner, F., Cochrane, A., & Green, C. S. (2017). Trial-dependent psychometric functions accounting for perceptual learning in 2-AFC discrimination tasks. *Journal of Vision*, 17, 1–16.

Katzner, S., Busse, L., & Carandini, M. (2011). GABAA inhibition controls response gain in visual cortex. *Journal of Neuroscience*, 31(16), 5931–5941.

Kaufman, L. & Pitblado, C. (1965). Further observations on the nature of effective binocular disparities. *The American journal of psychology*, 78(3), 379–391.

Kellman, P. J. (2013). Adaptive and Perceptual Learning Technologies in Medical Education and Training. *Military Medicine*, 178(10S), 98–106.

Kempter, R., Gerstner, W., & Van Hemmen, J. L. (1999). Hebbian learning and spiking neurons. *Polski tygodnik lekarski (Warsaw, Poland : 1960)*, 59(4), 4498.

Khalil, N. M., Nicotra, A., & Wilkins, A. J. (2011). Asymmetry of visual function in migraine with aura: correlation with lateralisation of headache and aura. *Cephalalgia*, 31(2), 213–21.

Kingdom, F. A. & Prins, N. (2010). *Psychophysics: a practical introduction*. Amsterdam: Elsevier.

Kingdom, F. A. & Prins, N. (2016). *Psychophysics: A Practical Introduction: Second Edition*. Elsevier Ltd., 2nd edition.

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., Broussard, C., & others (2007). What's new in Psychtoolbox-3. *Perception*, 36(14), 1.

Knoblauch, K. & Maloney, L. T. (2012). *Modeling Psychophysical Data in R*. New York, NY: Springer New York.

Konrad, J., Lacotte, B., & Dubois, E. (2000). Cancellation of image crosstalk in time-sequential displays of stereoscopic video. *IEEE Transactions on Image Processing*, 9(5), 897–908.

Kosslyn, S. M., Cacioppo, J. T., Davidson, R. J., Hugdahl, K., Lovallo, W. R., Spiegel, D., & Rose, R. (2002). Bridging psychology and biology: The analysis of individuals in groups. *American Psychologist*, 57(5), 341–351.

Kourtzi, Z., Betts, L. R., Sarkheil, P., & Welchman, A. E. (2005). Distributed neural plasticity for shape learning in the human visual cortex. *PLoS Biology*, 3(7), 1317–1327.

Krug, K., Brunskill, E., Scarna, A., Goodwin, G. M., & Parker, A. J. (2008). Perceptual switch rates with ambiguous structure-from-motion figures in bipolar disorder. *Proceedings of the Royal Society B: Biological Sciences*, 275(1645), 1839–1848.

Krug, K., Cumming, B. G., & Parker, A. J. (2004). Comparing perceptual signals of single V5/MT neurons in two binocular depth tasks. *Journal of neurophysiology*, 92(3), 1586–1596.

Kuai, S.-G., Zhang, J.-Y., Klein, S. A., Levi, D. M., & Yu, C. (2005). The essential role of stimulus temporal patterning in enabling perceptual learning. *Nature Neuroscience*, 8(11), 1497–1499.

Kundel, H. L., Nodine, C. F., Conant, E. F., & Weinstein, S. P. (2007). Holistic component of image perception in mammogram interpretation: gaze-tracking study. *Radiology*, 242(2), 396–402.

Lachaud, C. M. & Renaud, O. (2011). A tutorial for analyzing human reaction times: How to filter data, manage missing values, and choose a statistical model. *Applied Psycholinguistics*, 32(2), 389–416.

Laming, D. (2011). Fechner's Law: Where does the log transform come from? In J. A. Solomon (Ed.), *Fechner's Legacy in Psychology: 150 Years of Elementary Psychophysics* chapter 2, (pp. 7–23). Leiden: Brill, 1st edition.

Lamme, V. A. F. (2003). Recurrent corticocortical interactions in neural disease. *Archives of Neurology*, 60(2), 178–184.

Landy, M. S. & Johnston, E. B. (1995). Measurement and Modeling of Depth Cue. *Vision research*, 35(3), 389–412.

Langley, K., Fleet, D. J. D., & Hibbard, P. B. (1998). Linear and nonlinear transparencies in binocular vision. *Proceedings of the Royal Society of London B: Biological Sciences*, 265(1408), 1837–1845.

Langley, K., Fleet, D. J. D., & Hibbard, P. B. (1999). Stereopsis from contrast envelopes. *Vision Research*, 39(14), 2313–2324.

Law, C. & Gold, J. I. (2008). Neural correlates of perceptual learning in a sensory-motor, but not a sensory, cortical area. *Nature Neuroscience*, 11(4), 505–513.

Law, C. T. & Gold, J. I. (2009). Reinforcement learning can account for associative and perceptual learning on a visual-decision task. *Nature Neuroscience*, 12(5), 655–663.

Lawless, H. T. (2013). *Quantitative Sensory Analysis*. Chichester, UK: John Wiley & Sons.

Leek, M. R. (2001). Adaptive Procedure in Psychophysical Research. *Perception & Psychophysics*, 63(8), 1279–1292.

Lennie, P. (1998). Single units and visual cortical organization. *Perception*, 27, 889–935.

Levi, A., Shaked, D., Tadin, D., & Huxlin, K. R. (2015). Is improved contrast sensitivity a natural consequence of visual training? *Journal of Vision*, 14(10), 1158–1158.

Levi, D. M. (2005). Perceptual learning in adults with amblyopia: A reevaluation of critical periods in human vision. *Developmental Psychobiology*, 46(3), 222–232.

Levi, D. M., Klein, S. A., & Chen, I. (2005). What is the signal in noise? *Vision Research*, 45(14), 1835–1846.

Levi, D. M. & Li, R. W. (2009). Perceptual learning as a potential treatment for amblyopia: A mini-review. *Vision Research*, 49(21), 2535–2549.

Levi, D. M. & Polat, U. (1996). Neural plasticity in adults with amblyopia. *Investigative Ophthalmology and Visual Science*, 37(3), 6830–6834.

Levi, D. M., Polat, U., & Hu, Y. S. (1997). Improvement in Vernier acuity in adults with amblyopia: Practice makes better. *Investigative Ophthalmology and Visual Science*, 38(8), 1493–1510.

Levitt, H. (1971). Transformed Up-Down Methods in Psychoacoustics. *The Journal of the Acoustical Society of America*, 49(2B), 467–477.

Lewis, T. L. & Maurer, D. (2005). Multiple sensitive periods in human visual development: Evidence from visually deprived children. *Developmental Psychobiology*, 46(3), 163–183.

Li, P., Zhu, S., Chen, M., Han, C., Xu, H., Hu, J., Fang, Y., & Lu, H. D. (2013). A Motion direction preference map in monkey V4. *Neuron*, 78(2), 376–388.

Lin, Z., Dosher, B. A., & Lu, Z.-L. (2017). Mixture of easy trials enables transient and sustained perceptual improvements through priming and perceptual learning. *Scientific Reports*, 7(1), 7421.

Lin, Z. & Murray, S. O. (2014). Priming of awareness or how not to measure visual awareness. *Journal of Vision*, 14(2014), 1–17.

Lin, Z. & Murray, S. O. (2015). Automaticity of Unconscious Response Inhibition : Comment on Chiu and Aron (2014). *Journal of Experimental Psychology: General*, 144(1), 244–254.

Ling, S. & Carrasco, M. (2006). Sustained and transient covert attention enhance the signal via different contrast response functions. *Vision Research*, 46(8-9), 1210–1220.

Lippert, J. J. & Wagner, H. (2001). A threshold explains modulation of neural responses to opposite-contrast stereograms. *Neuroreport*, 12(15), 3205–3208.

Liu, J., Dosher, B. A., & Lu, Z. L. (2014). Modeling trial by trial and block feedback in perceptual learning. *Vision Research*, 99, 46–56.

Liu, J., Lu, Z.-L., & Dosher, B. A. (2010). Augmented Hebbian reweighting: interactions between feedback and training accuracy in perceptual learning. *Journal of vision*, 10(10), 29.

Liu, J., Lu, Z. L., & Dosher, B. A. (2012a). Mixed training at high and low accuracy levels leads to perceptual learning without feedback. *Vision Research*, 61, 15–24.

Liu, L., Wang, F., Zhou, K., Ding, N., & Luo, H. (2017). Perceptual integration rapidly activates dorsal visual pathway to guide local processing in early visual areas. *PLoS Biology*, 15(11), 1–15.

Liu, M. M., Chan, C.-C., & Tuo, J. (2012b). Genetic mechanisms and age-related macular degeneration: common variants, rare variants, copy number variations, epigenetics, and mitochondrial genetics. *Human Genomics*, 6(1), 13.

Liu, P., Forte, J., Sewell, D., & Carter, O. (2018). Cognitive load effects on early visual perceptual processing. *Attention, Perception, & Psychophysics*, 80(4), 929–950.

Liu, Z. (1999). Perceptual learning in motion discrimination that generalizes across motion directions. *Proceedings of the National Academy of Sciences of the United States of America*, 96(24), 14085–14087.

Liu, Z. & Weinshall, D. (2000). Mechanisms of generalization in perceptual learning. *Vision Research*, 40(1), 97–109.

Livingstone, M. & Hubel, D. H. (1988). Segregation of Form, Color, Movement, and Depth: Anatomy, Physiology, and Perception. *Science*, 240(4835), 740–749.

Livingstone, M. S. & Hubel, D. H. (1987). Psychophysical of Form, Color, Movement, and Depth. *Journal of Neuroscience*, 7(November).

Lu, Z.-L. & Dosher, B. A. (2004). Perceptual learning retunes the perceptual template in foveal orientation identification. *Journal of Vision*, 4(1), 5–5.

Lu, Z.-L., Dosher, B. A., & Lu, Zhong Lin, B. D. (2014). *Visual psychophysics: From laboratory to theory*, volume 43. MIT Press.

Lu, Z. L., Hua, T., Huang, C. B., Zhou, Y., & Dosher, B. A. (2011). Visual perceptual learning. *Neurobiology of learning and memory*, 95(2), 145–51.

Lu, Z.-L., Lesmes, L. A., & Sperling, G. (1999). The mechanism of isoluminant chromatic motion perception. *Proceedings of the National Academy of Sciences*, 96(14), 8289–8294.

Lui, L. L., Bourne, J. A., & Rosa, M. G. P. (2007). Spatial and temporal frequency selectivity of neurons in the middle temporal visual area of new world monkeys (Callithrix jacchus). *European Journal of Neuroscience*, 25(6), 1780–1792.

Lunghi, C. & Sale, A. (2015). A cycling lane for brain rewiring. *Current Biology*, 25(23), R1122–R1123.

Maniglia, M., Pavan, A., Sato, G., Contemori, G., Montemurro, S., Battaglini, L., & Casco, C. (2016). Perceptual learning leads to long lasting visual improvement in patients with central vision loss. *Restorative Neurology and Neuroscience*, 34(5), 697–720.

Maniglia, M. & Seitz, A. R. (2018). Towards a whole brain model of Perceptual Learning. *Current Opinion in Behavioral Sciences*, 20, 47–55.

Maniglia, M., Soler, V., Cottereau, B., & Trotter, Y. (2018). Spontaneous and training-induced cortical plasticity in MD patients: Hints from lateral masking. *Scientific Reports*, 8(1), 1–11.

Manning, C., Jones, P. R., Dekker, T. M., & Pellicano, E. (2018). Psychophysics with children: Investigating the effects of attentional lapses on threshold estimates. *Attention, Perception, and Psychophysics*, 80(5), 1311–1324.

Manning, C., Tibber, M. S., Charman, T., Dakin, S. C., & Pellicano, E. (2015). Enhanced Integration of Motion Information in Children With Autism. *Journal of Neuroscience*, 35(18), 6979–6986.

Mansouri, B. & Hess, R. F. (2006). The global processing deficit in amblyopia involves noise segregation. *Vision Research*, 46(24), 4104–4117.

Markov, N. T., Vezoli, J., Chameau, P., Falchier, A., Quilodran, R., Huissoud, C., Lamy, C., Misery, P., Giroud, P., Ullman, S., Barone, P., Dehay, C., Knoblauch, K., & Kennedy, H. (2014). Anatomy of hierarchy: Feedforward and feedback pathways in macaque visual cortex. *Journal of Comparative Neurology*, 522(1), 225–259.

Martinez, L. M. & Alonso, J.-M. (2003). Complex receptive fields in primary visual cortex. *The neuroscientist*, 9(5), 317–331.

Mather, G., Verstraten, F., & Anstis, S. M. (1998). *The motion aftereffect: A modern perspective*. Mit Press.

Maunsell, J., Nealey, T., & DePriest, D. (1990). Magnocellular and parvocellular contributions to responses in the middle temporal visual area (MT) of the macaque monkey. *Journal of Neuroscience*, 10(October), 3323–3334.

McColl, S. L. & Wilkinson, F. (2000). Visual contrast gain control in migraine: Measures of visual cortical excitability and inhibition. *Cephalalgia*, 20(2), 74.

McColl, S. L., Ziegler, L., & Hess, R. F. (2000). Stereodeficient subjects demonstrate non-linear stereopsis. *Vision Research*, 40(9), 1167–1177.

McGovern, D. P., Webb, B. S., & Peirce, J. W. (2012). Transfer of perceptual learning between different visual tasks. *Journal of Vision*, 12(11), 4–4.

McKee, S. P., Verghese, P., & Farell, B. (2004). What is the depth of a sinusoidal grating? *Journal of Vision*, 4(7), 1.

Mckee, S. P. & Westheimer, G. (1978). Improvement in vernier acuity with practice. *Perception & psychophysics*, 24(3), 258–262.

Mckendrick, A. & Badcock, D. (2003). Changes in visual field performance after migraine measured with flickering stimuli. *Cephalalgia*, 23(7), 743–744.

McKendrick, A. M., Badcock, D. R., & Gurgone, M. (2006). Vernier acuity is normal in migraine, whereas global form and global motion perception are not. *Investigative Ophthalmology and Visual Science*, 47(7), 3213–3219.

McKendrick, A. M., Chan, Y. M., Vingrys, A. J., Turpin, A., & Badcock, D. R. (2017). Daily vision testing can expose the prodromal phase of migraine. *Cephalalgia*, (pp. 0333102417741130).

McKendrick, A. M. & Sampson, G. P. (2009). Low spatial frequency contrast sensitivity deficits in migraine are not visual pathway selective. *Cephalalgia*, 29(5), 539–549.

McKendrick, A. M., Vingrys, A. J., Badcock, D. R., & Heywood, J. T. (2001). Visual dysfunction between migraine events. *Investigative Ophthalmology and Visual Science*, 42(3), 626–633.

Mclaren, P. L. & Mackintosh, N. J. (2000). An elemental model of associative learning I. Latent inhibition and perceptual learning. *Animal Learning & Behavior*, 28(3), 211–246.

Mendes, L. C., Lúcia, M., Simas, D. B., Galdino, M. K. C., Vieira, J. G., Simas, M. L. d. B., & Santos, N. A. d. (2011). Evaluation of contrast sensitivity among patients with migraine. *Psicologia USP*, 22(1), 81–97.

Merigan, W., Katz, L., & Maunsell, J. (1991). The effects of parvocellular lateral geniculate lesions on the acuity and contrast sensitivity of macaque monkeys. *The Journal of Neuroscience*, 11(4), 994–1001.

Mikami, A., Newsome, W. T., & Wurtz, R. H. (1986). Motion selectivity in macaque visual cortex. II. Spatiotemporal range of directional interactions in MT and V1. *Journal of neurophysiology*, 55(6), 1328–39.

Miłkowski, M. (2018). Mechanisms and the mental. In S. Glennan & P. Illaria (Eds.), *The Routledge Handbook of Mecahnisms and Mechanical Philosophy* chapter 6, (pp. 74–88). Oxon: Routledge, 1st edition.

Milner, B., Squire, L. R., & Kandel, E. R. (1998). Cognitive neuroscience and the study of memory. *Neuron*, 20(3), 445–468.

Mollon, J. D., Bosten, J. M., Peterzell, D. H., & Webster, M. A. (2017). Individual differences in visual science: What can be learned and what is good experimental practice? *Vision Research*, 141(October), 4–15.

Mollon, J. D. & Danilova, M. V. (1996). Three remarks on perceptual learning. *Spatial vision*, 10(1), 51–58.

Moscatelli, A., Mezzetti, M., & Lacquaniti, F. (2012). Modeling psychophysical data at the population-level: The generalized linear mixed model. *Journal of Vision*, 12(11), 26–26.

Movshon, J. A. & Lennie, P. (1979). Pattern-selective adaptation in visual cortical neurones. *Nature*, 278(5707), 850–852.

Movshon, J. A. & Newsome, W. T. (1996). Visual response properties of striate cortical neurons projecting to area MT in macaque monkeys. *Journal of Neuroscience*, 16(23), 7733–7741.

Movshon, J. A., Thompson, I. D., & Tolhurst, D. J. (1978). Spatial and temporal contrast sensitivity of neurones in areas 17 and 18 of the cat visual cortex. *J. Physiol.*, 283, 101–120.

Muckli, L., Kohler, A., Kriegeskorte, N., & Singer, W. (2005). Primary visual cortex activity along the apparent-motion trace reflects illusory perception. *PLoS Biology*, 3(8).

Nakayama, K. & Shimojo, S. (1990). Da Vinci stereopsis: Depth and subjective occluding contours from unpaired image points. *Vision research*, 30(11), 1811–1825.

Nankoo, J. F., Madan, C. R., Spetch, M. L., & Wylie, D. R. (2015). Temporal summation of global form signals in dynamic Glass patterns. *Vision Research*, 107, 30–35.

Nealey, S. M. & Riley, D. A. (1963). Loss and Recovery of Discrimination of Visual Depth in Dark-Reared Rats. *The American Journal of Psychology*, 76(2), 329–332.

Nefs, H. T., O'Hare, L., & Harris, J. M. (2010). Two independent mechanisms for motion-in-depth perception: Evidence from individual differences. *Frontiers in Psychology*, 1(OCT), 1–8.

Nelles, G., Esser, J., Eckstein, A., Tiede, A., Gerhard, H., & Diener, H. C. (2001). Compensatory visual field training for patients with hemianopia after stroke. *Neuroscience Letters*, 306(3), 189–192.

Neri, P. (2004). Stereoscopic Processing of Absolute and Relative Disparity in Human Visual Cortex. *Journal of Neurophysiology*, 92(3), 1880–1891.

Nessler, B., Pfeiffer, M., & Maass, W. (2008). Hebbian Learning of Bayes Optimal Decisions. *Advances in Neural Information Processing Systems*, (pp. 1169–1176).

Neves, G., Cooke, S. F., & Bliss, T. V. (2008). Synaptic plasticity, memory and the hippocampus: A neural network approach to causality. *Nature Reviews Neuroscience*, 9(1), 65–75.

Neville, H. J. & Bavelier, D. (1998). Neural organization and plasticity of language. *Current opinion in Neurobiology*, 8(2), 254–258.

Newsome, W. T., Britten, K. H., & Movshon, J. A. (1989). Neuronal correlates of a perceptual decision. *Nature*, 341(6237), 52–54.

Newsome, W. T., Park, B., York, N., & Brook, S. (1988). A Selective Impairment of Motion Perception Following Lesions of the Middle Temporal Visual Area (MT). *The Journal of Neuroscience*, 8(6).

Nieder, A. & Wagner, H. (2001). Hierarchical processing of horizontal disparity information in the visual forebrain of behaving owls. *Journal of Neuroscience*, 21(12), 4514–4522.

Nienborg, H., Bridge, H., Parker, A. J., & Cumming, B. G. (2004). Receptive field size in V1 neurons limits acuity for perceiving disparity modulation. *Journal of Neuroscience*, 24(9), 2065–2076.

Nishida, S. (2011). Advancement of motion psychophysics : Review 2001 – 2010. *Science*, 11(2011), 1–53.

Nityananda, V., Tarawneh, G., Henriksen, S., Umeton, D., Simmons, A., & Read, J. C. (2018). A Novel Form of Stereo Vision in the Praying Mantis. *Current Biology*, 28(4), 588–593.

O'Hare, L. & Hibbard, P. B. (2016). Visual processing in migraine. *Cephalalgia*, 36(11), 1057–1076.

Ohzawa, I., DeAngelis, G. C., & Freeman, R. D. (1990). Stereoscopic depth discrimination in the visual cortex: Neurons ideally suited as disparity detectors. *Science*, 249(4972), 1037–1041.

Ohzawa, I., DeAngelis, G. C., & Freeman, R. D. (1997). Encoding of binocular disparity by complex cells in the cat's visual cortex. *Journal of neurophysiology*, 77(6), 2879–2909.

Orban, G. A., Janssen, P., & Vogels, R. (2006). Extracting 3D structure from disparity. *Trends in neurosciences*, 29(8), 466–473.

Oruç, I., Maloney, L. T., & Landy, M. S. (2003). Weighted linear cue combination with possibly correlated error. *Vision research*, 43(23), 2451–2468.

Ostwald, D., Lam, J. M., Li, S., & Kourtzi, Z. (2008). Neural Coding of Global Form in the Human Visual Cortex. *Journal of Neurophysiology*, 99(5), 2456–2469.

Owsley, C. (2011). Aging and vision. *Vision Research*, 51(13), 1610–1622.

Palmer, J. E., Chronicle, E. P., Rolan, P., & Mulleners, W. M. (2000). Cortical hyperexcitability is cortical under-inhibition: Evidence from a novel functional test of migraine patients. *Cephalalgia*, 20(6), 525–532.

Parker, A. J. (2007). Binocular depth perception and the cerebral cortex. *Nature Reviews Neuroscience*, 8(5), 379–91.

Pascual-Leone, A. & Walsh, V. (2001). Fast Backprojections from the Motion to the Primary Visual Area Necessary for Visual Awareness. *Science*, 292(5516), 510–512.

Pasternak, T. & Merigan, W. H. (1994). Motion Perception following Lesions of the Superior Temporal Sulcus in the Monkey. *Cerebral Cortex*, May/June(4), 247–259.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial vision*, 10(4), 437–442.

Peterzell, D., Read, J., & Serrano-Pedraza;, I. (2018). Individual differences: On the possible relativity of spatial-frequency-tuned stereoscopic processes underlying disparity threshold functions. *Journal of Vision*, 18(10), 134–134.

Petrov, A. A., Dosher, B. A., & Lu, Z.-L. (2005). The dynamics of perceptual learning: An incremental reweighting model. *Psychological Review*, 112(4), 715–743.

Petrov, A. A., Dosher, B. A., & Lu, Z. L. (2006a). Perceptual learning without feedback in non-stationary contexts: Data and model. *Vision Research*, 46(19), 3177–3197.

Petrov, Y., Verghese, P., & McKee, S. P. (2006b). Collinear facilitation is largely uncertainty reduction. *Journal of vision*, 6(2), 170–8.

Plank, T., Rosengarth, K., Schmalhofer, C., Goldhacker, M., Brandl-Rühle, S., & Greenlee, M. W. (2014). Perceptual learning in patients with macular degeneration. *Frontiers in Psychology*, 5(OCT), 1–14.

Poggio, T., Fahle, M., & Edelman, S. (1991). Fast perceptual learning in visual hyperacuity. *Massachusetts Inst of Tech Artfical Intelligence Lab*.

Polat, U. (2009). Making perceptual learning practical to improve visual functions. *Vision Research*, 49(21), 2566–2573.

Polat, U., Ma-Naim, T., Belkin, M., & Sagi, D. (2004). Improving vision in adult amblyopia by perceptual learning. *Proceedings of the National Academy of Sciences*, 101(17), 6692–6697.

Polat, U. & Sagi, D. (1993). Lateral Interactions Between Spatial Channels : Suppression and Facilitation Revealed by Lateral Masking Experiments. *Vision Research*, 33(7), 993–999.

Prince, S., Cumming, B. G., & Parker, A. J. (2017). Range and Mechanism of Encoding of Horizontal Disparity in Macaque V1. *Journal of Neurophysiology*, 87(1), 209–221.

Prins, N. & Kingdom, F. A. (2009). Palamedes: Matlab routines for analyzing psychophysical data.

Prins, N. & Kingdom, F. A. (2018). Applying the model-comparison approach to test specific research hypotheses in psychophysical research using the Palamedes toolbox. *Frontiers in Psychology*, 9(JUL), 1–14.

Queiroz, L. P., Rapoport, A. M., Weeks, R. E., Sheftell, F. D., Siegel, S. E., & Baskin, S. M. (1997). Characteristics of migraine visual aura. *Headache*, 37(3), 137–141.

Ramachandran, V. S. & Braddick, O. J. (1973). Orientation specific learning in stereopsis. *Perception*, 2(3), 371–376.

Read, J. C., Parker, A. J., & Cumming, B. G. (2002). A simple model accounts for the response of disparity-tuned V1 neurons to anticorrelated images. *Visual Neuroscience*, 19(6), 735–753.

Read, J. C. A. & Eagle, R. A. (2000). Reversed stereo depth and motion direction with anti-correlated stimuli. *Vision research*, 40(24), 3345–3358.

Rees, G., Friston, K., & Koch, C. (2000). A direct quantitative relationship between the functional properties of human and macaque V5. *Nature Neuroscience*, 3(7), 716–723.

Regan, D., Erkelens, C. J., & Collewijn, H. (1986). Necessary conditions for the perception of motion in depth. *Investigative Ophthalmology & Visual Science*, 27(4), 584–597.

Reisen, A. H. (1947). The development of visual perception in man and chimpanzee. *Science*, 106(Table 1), 107–108.

Ren, Z., Zhou, J., Yao, Z., Wang, Z., Yuan, N., Xu, G., Wang, X., Zhang, B., Hess, R. F., & Zhou, Y. (2016). Neuronal basis of perceptual learning in striate cortex. *Nature Scientific Reports*, 6, 24769.

Rescorla, R. (2008). {R}escorla-{W}agner model. *Scholarpedia*, 3(3), 2237.

Rescorla, R. A. & Wagner, A. R. (1972). A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and non reinforcement. In B. A.H. & P. W.F. (Eds.), *Classical conditioning II: current research and theory* (pp. 1–18). New York, NY: Appleton-Century-Crofts.

Rockland, K. S. & Knutson, T. (2000). Feedback Connections From Area Mt Of The Squirrel-Monkey To Areas V1 And V2. *Journal Of Comparative Neurology*, 425(3), 345–368.

Roe, A. W., Chelazzi, L., Connor, C. E., Conway, B. R., Fujita, I., Gallant, J. L., Lu, H., & Vanduffel, W. (2012). Toward a Unified Theory of Visual Area V4. *Neuron*, 74(1), 12–29.

Rogers, B. J. & Anstis, S. M. (1975). Reversed depth from positive and negative stereograms. *Perception*, 4(2), 193–201.

Romei, V., Chiappini, E., Hibbard, P. B., & Avenanti, A. (2016a). Empowering Reentrant Projections from V5 to V1 Boosts Sensitivity to Motion. *Current Biology*, 26(16), 2155–2160.

Romei, V., Thut, G., & Silvanto, J. (2016b). Information-Based Approaches of Noninvasive Transcranial Brain Stimulation. *Trends in Neurosciences*, 39(11), 782–795.

Rudolph, K. & Pasternak, T. (1999). Transient and permanent deficits in motion perception after lesions of cortical areas MT and MST in the macaque monkey. *Cerebral Cortex*, 9(1), 90–100.

Saarinen, J. & Levi, D. M. (1995). Perceptual learning in vernier acuity: What is learned? *Vision Research*, 35(4), 519–527.

Sagi, D. (2011). Perceptual learning in Vision Research. *Vision Research*, 51(13), 1552–1566.

Sahraie, A., Trevethan, C. T., MacLeod, M. J., Murray, A. D., Olson, J. A., & Weiskrantz, L. (2006). Increased sensitivity after repeated stimulation of residual spatial channels in blindsight. *Proceedings of the National Academy of Sciences of the United States of America*, 103(40), 14971–14976.

Sasaki, Y., Nanez, J. E., & Watanabe, T. (2010). Advances in visual perceptual learning and plasticity. *Nat Rev Neurosci*, 11(1), 53–60.

Scarfe, P. & Glennerster, A. (2014). Humans Use Predictive Kinematic Models to Calibrate Visual Cues to Three-Dimensional Surface Slant. *Journal of Neuroscience*, 34(31), 10394–10401.

Scarfe, P. & Glennerster, A. (2015). Using high-fidelity virtual reality to study perception in freely moving observers. *Journal of Vision*, 15(9), 3.

Schiller, P. H., Logothetis, N. K., & Charles, E. R. (1991). Parallel pathways in the visual system: Their role in perception at isoluminance. *Neuropsychologia*, 29(6), 433–441.

Schmitt, C., Kromeier, M., Bach, M., & Kommerell, G. (2002). Interindividual variability of learning in stereoacuity. *Graefe's Archive for Clinical and Experimental Ophthalmology*, 240(9), 704–709.

Schor, C. M., Edwards, M., & Pope, D. R. (1998). Spatial-frequency and contrast tuning of the transient-stereopsis system. *Vision research*, 38(20), 3057–3068.

Schoups, A. A., Vogels, R., & Orban, G. A. (1995). Human perceptual learning in identifying the oblique orientation: retinotopy, orientation specificity and monocularity. *The Journal of physiology*, 483.3, 797–810.

Schoups, A. A., Vogels, R., Qian, N., & Orban, G. (2001). Practising orientation identification improves orientation coding in V1 neurons. *Nature*, 412(6846), 549–553.

Schwartz, S., Maquet, P., & Frith, C. (2002). Neural correlates of perceptual learning: a functional MRI study of visual texture discrimination. *Proceedings of the National Academy of Sciences of the United States of America*, 99(26), 17137–17142.

Seitz, A. R. (2017). Perceptual learning. *Current Biology*, 27(13), R631–R636.

Seitz, A. R. & Dinse, H. R. (2007). A common framework for perceptual learning. *Current Opinion in Neurobiology*, 17(2), 148–153.

Seitz, A. R., Nanez, J. E., Holloway, S., Tsushima, Y., & Watanabe, T. (2006). Two cases requiring external reinforcement in perceptual learning. *Journal of Vision*, 6(9), 9.

Seitz, A. R., Nanez, J. E., Holloway, S. R., Koyama, S., & Watanabe, T. (2005). Seeing what is not there shows the costs of perceptual learning. *Proceedings of the National Academy of Sciences of the United States of America*, 102(25), 9080–9085.

Seitz, A. R. & Watanabe, T. (2005). A unified model for perceptual learning. *Trends in Cognitive Sciences*, 9(7), 329–334.

Seitz, A. R. & Watanabe, T. (2009). The phenomenon of task-irrelevant perceptual learning. *Vision Research*, 49(21), 2604–2610.

Sereno, M. I. (2005). Plasticity and its limits. *Nature*, 435(7040), 288–9.

Seriès, P., Latham, P. E., & Pouget, A. (2004). Tuning curve sharpening for orientation selectivity: coding efficiency and the impact of correlations. *Nature Neuroscience*, 7(10), 1129–1135.

Shen, Y. (2013). Comparing adaptive procedures for estimating the psychometric function for an auditory gap detection task. *Attention, Perception, and Psychophysics*, 75(4), 771–780.

Shepherd, A. J. (2000). Visual contrast processing in migraine. *Cephalalgia*, 20(10), 865–880.

Shepherd, A. J. (2001). Increased visual after-effects following pattern adaptation in migraine: a lack of intracortical excitation? *Brain : a journal of neurology*, 124(Pt 11), 2310–2318.

Shepherd, A. J. (2006). Local and global motion after-effects are both enhanced in migraine, and the underlying mechanisms differ across cortical areas. *Brain*, 129(7), 1833–1843.

Shepherd, A. J., Beaumont, H. M., Hine, T. J., & Street, M. (2012). Motion processing deficits in migraine are related to contrast sensitivity. *Cephalalgia*, 32(7), 554–570.

Shepherd, A. J., Hine, T. J., & Beaumont, H. M. (2013). Color and spatial frequency are related to visual pattern sensitivity in migraine. *Headache*, 53(7), 1087–1103.

Shepherd, A. J., Wyatt, G., & Tibber, M. S. (2011). Visual metacontrast masking in migraine. *Cephalalgia : an international journal of headache*, 31(3), 346–56.

Shergill, S. S., Samson, G., Bays, P. M., Frith, C. D., & Wolpert, D. M. (2005). Evidence for sensory prediction deficits in schizophrenia. *American Journal of Psychiatry*, 162(12), 2384–2386.

Shibata, K., Sagi, D., & Watanabe, T. (2014). Two-stage model in perceptual learning: Toward a unified theory. *Annals of the New York Academy of Sciences*, 1316(1), 18–28.

Shibata, K., Watanabe, T., Sasaki, Y., & Kawato, M. (2011). Perceptual learning incepted by decoded fMRI neurofeedback without stimulus presentation. *Science (New York, N.Y.)*, 334(6061), 1413–5.

Shibata, K., Yamagishi, N., Ishii, S., & Kawato, M. (2009). Boosting perceptual learning by fake feedback. *Vision Research*, 49(21), 2574–2585.

Shiu, L. P. & Pashler, H. (1992). Improvement in line orientation discrimination is retinally local but dependent on cognitive set. *Perception & psychophysics*, 52(5), 582–588.

Sillito, A. M., Cudeiro, J., & Jones, H. E. (2006). Always returning: feedback and sensory processing in visual cortex and thalamus. *Trends in Neurosciences*, 29(6), 307–316.

Simmons, D. R., Robertson, A. E., McKay, L. S., Toal, E., McAleer, P., & Pollick, F. E. (2009). Vision in autism spectrum disorders. *Vision Research*, 49(22), 2705–2739.

Simoncelli, E. P. (2003). Vision and the statistics of the visual environment. *Current Opinion in Neurobiology*, 13(2), 144–149.

Simoncelli, E. P. & Heeger, D. J. (1998). A Model of Neuronal Responses in Visual Area MT. *Vision Research*, 38(5), 743–761.

Sincich, L. C., Adams, D. L., & Horton, J. C. (2003). Complete flatmounting of the macaque cerebral cortex. *Visual Neuroscience*, 20(6), 663–686.

Sincich, L. C. & Horton, J. C. (2005). The Circutrly of V1 and V2: Integration of Color, Form, and Motion. *Annual Review of Neuroscience*, 28(1), 303–326.

Singh, P. & Shepherd, A. J. (2016). Enhanced Motion Aftereffects in Migraine Are Related to Contrast Sensitivity: Implications for Models of Differences in Precortical/Cortical FunctionEnhanced Motion Aftereffects in Migraine. *Investigative ophthalmology & visual science*, 57(3), 1228–1234.

Smallman, H. S. & MacLeod, D. I. A. (1997). Spatial scale interactions in stereo sensitivity and the neural representation of binocular disparity. *Perception*, 26(8), 977–994.

Smirnakis, S. M., Brewer, A. A., Schmid, M. C., Tolias, A. S., Schüz, A., Augath, M., Inhoffen, W., Wandell, B. A., & Logothetis, N. K. (2005). Lack of long-term cortical reorganization after macaque retinal lesions. *Nature*, 435(7040), 300–307.

Smith, a. T., Snowden, R. J., & Milne, a. B. (1994). Is global motion really based on spatial integration of local motion signals? *Vision research*, 34(18), 2425–2430.

Solgi, M., Liu, T., & Weng, J. (2013). A computational developmental model for specificity and transfer in perceptual learning. *Journal of Vision*, 13, 7.

Solomon, J. A. & Tyl (2017). The improvement of contrast sensitivity with practice is not compatible with a sensory threshold account. *Journal of the Optical Society of America A*, 34(6), 1–11.

Song, S., Miller, K. D., & Abbott, L. F. (2000). Competitive Hebbian learning through spike-timing-dependent synaptic plasticity. *Nature neuroscience*, 3(9), 919–926.

Sowden, P. T., Davies, I. R. L., & Roling, P. (2000). Perceptual Learning of the Detection of Features in X-Ray Images : A Functional Role for Improvements in Adults ' Visual Sensitivity ? *Journal of Experimental Psychology*, 26(1), 379–390.

Sowden, P. T., Rose, D., & Davies, I. R. L. (2002). Perceptual learning of luminance contrast detection: Specific for spatial frequency and retinal location but not orientation. *Vision Research*, 42(10), 1249–1258.

Sterkin, A., Levy, Y., Pokroy, R., Lev, M., Levian, L., Doron, R., Yehezkel, O., Fried, M., Frenkel-Nir, Y., Gordon, B., & Polat, U. (2018). Vision improvement in pilots with presbyopia following perceptual learning. *Vision Research*, 152(November), 61–73.

Stevens, S. S. (1957). On the psychophysical law. *Psychological Review*, 64(3), 153–181.

Stinson, C. & Sullivan, J. (2018). Mechanistic explanation in neuroscience. In S. Glennan & P. Illaria (Eds.), *The Routledge Handbook of Mecahnisms and Mechanical Philosophy* chapter 28, (pp. 375–387). Oxon: Routledge, 1st edition.

Stovner, L. J. & Andree, C. (2010). Prevalence of headache in Europe: a review for the Eurolight project. *The journal of headache and pain*, 11(4), 289.

Sundareswaran, V. & Vaina, L. M. (1995). Learning direction in global motion: two classes of psychophysically-motivated models. *Advances in Neural Information Processing Systems 7*, (pp. 917–924).

Sutton, R. S. & Barto, A. G. (2011). *Reinforcement Learning: An Introduction*. MIT press, 2nd edition.

Swanson, W. H. & Birch, E. E. (1992). Extracting thresholds from noisy psychophysical data. *Perception & Psychophysics*, 51(5), 409–422.

Talluri, B. C., Hung, S.-C., Seitz, A. R., & Seriès, P. (2015). Confidence-based integrated reweighting model of task-difficulty explains location-based specificity in perceptual learning. *Journal of vision*, 15(10), 17.

Tanabe, S., Umeda, K., & Fujita, I. (2004). Rejection of false matches for binocular correspondence in macaque visual cortical area V4. *Journal of Neuroscience*, 24(37), 8170–8180.

Tanabe, S., Yasuoka, S., & Fujita, I. (2008). Disparity-energy signals in perceived stereoscopic depth. *Journal of vision*, 8(3), 22.

Tanaka, H. & Ohzawa, I. (2006). Neural basis for stereopsis from second-order contrast cues. *Journal of Neuroscience*, 26(16), 4370–4382.

Thiele, J. (2012). Potential of GLMM in modelling invasive spread. *CAB Reviews: Perspectives in Agriculture, Veterinary Science, Nutrition and Natural Resources*, 7(016), 1–10.

Thomas, O. M., Cumming, B. G., & Parker, A. J. (2002). A specialization for relative disparity in V2. *Nature neuroscience*, 5(5), 472.

Tibber, M. S., Guedes, A., & Shepherd, A. J. (2006). Orientation discrimination and contrast detection thresholds in migraine for cardinal and oblique angles. *Investigative Ophthalmology and Visual Science*, 47(12), 5599–5604.

Tibber, M. S., Kelly, M. G., Jansari, A., Dakin, S. C., & Shepherd, A. J. (2014). An inability to exclude visual noise in migraine. *Investigative Ophthalmology and Visual Science*, 55(4), 2539–2546.

Tong, F. (2018). Foundations of Vision. In J. T. Wixted & J. T. Serences (Eds.), *Steven's Handbook of Experimental Psychology and Cognitive Neuroscience. Volume 2. Sensation, Perception & Attention* chapter 1, (pp. 1–62). John Wiley & Sons, 4th edition.

Treutwein, B. (1995). Adaptive psychophysical procedures. *Vision Research*, 35(17), 2503–2522.

Treutwein, B. & Strasburger, H. (1999). Fitting the psychometric function. *Perception and Psychophysics*, 61(1), 87–106.

Trevethan, C. T., Urquhart, J., Ward, R., Gentleman, D., & Sahraie, A. (2012). Evidence for perceptual learning with repeated stimulation after partial and total cortical blindness. *Advances in Cognitive Psychology*, 8(1), 29–37.

Tsirlin, I., Wilcox, L. M., & Allison, R. S. (2010). Monocular occlusions determine the perceived shape and depth of occluding surfaces. *Journal of Vision*, 10(6), 11.

Tsirlin, I., Wilcox, L. M., & Allison, R. S. (2014). A computational theory of da Vinci stereopsis. *Journal of Vision*, 14(7), 5.

Tsodyks, M., Adini, Y., & Sagi, D. (2004). Associative learning in early vision. *Neural Networks*, 17(5-6), 823–832.

Tsodyks, M. & Gilbert, C. D. (2004). Neural Networks of Perceptual learning. *Nature*, 431(October), 775–781.

Umeda, K., Tanabe, S., & Fujita, I. (2007). Representation of stereoscopic depth based on relative disparity in macaque area V4. *Journal of neurophysiology*, 98(1), 241–252.

Urbanski, M., Coubard, O. A., & Bourlon, C. (2014). Visualizing the blind brain: brain imaging of visual field defects from early recovery to rehabilitation techniques. *Frontiers in Integrative Neuroscience*, 8(September), 74.

Vaina, L. M., Belliveau, J. W., Roziers, E. B. d., & Zeffiro, T. A. (1998). Neural systems underlying learning and representation of global motion. *Proceedings of the National Academy of Sciences*, 95(21), 12657–12662.

Vaina, L. M., Rana, K. D., Cotos, I., Li-Yang, C., Huang, M. A., & Podea, D. (2014). When does Subliminal Affective Image Priming Influence the Ability of Schizophrenic Patients to Perceive Face Emotions? *Medical Science Monitor*, 20, 2788–2798.

Vaina, L. M., Sundareswaran, V., & Harris, J. G. (1995). Learning to ignore: psychophysics and computational modeling of fast learning of direction in noisy motion stimuli. *Cognitive Brain Research*, 2(3), 155–163.

Wade, N. & Swanston, M. (2013). The Heritage. In *Visual perception: An introduction* chapter 2, (pp. 32–84). Psychology Press, 2nd edition.

Wagner, D., Manahilov, V., Gordon, G. E., & Loffler, G. (2013). Global shape processing deficits are amplified by temporal masking in migraine. *Investigative Ophthalmology and Visual Science*, 54(2), 1160–1168.

Wagner, D., Manahilov, V., Loffler, G., Gordon, G. E., & Dutton, G. N. (2010). Visual noise selectively degrades vision in migraine. *Investigative Ophthalmology and Visual Science*, 51(4), 2294–2299.

Walk, R. D., Trychin, S., & Karmel, B. Z. (1965). Depth perception in the dark-reared rat as a function of time in the dark. *Psychonomic Science*, 3, 9–10.

Wandell, B. A. & Smirnakis, S. M. (2009). Plasticity and stability of visual field maps in adult primary visual cortex. *Nature Neuroscience*, 10(12), 873–884.

Wang, R., Zhang, J.-Y., Klein, S. A., Levi, D. M., & Yu, C. (2014). Vernier perceptual learning transfers to completely untrained retinal locations after double training: A "piggybacking" effect. *Journal of Vision*, 14(13), 12–12.

Watanabe, T., Náñez, J. E., Koyama, S., Mukai, I., Liederman, J., & Sasaki, Y. (2002). Greater plasticity in lower-level than higher-level visual motion processing in a passive perceptual learning task. *Nature neuroscience*, 5(10), 1003–1009.

Watanabe, T., Náñez, J. E., & Sasaki, Y. (2001). Perceptual learning without perception. *Nature*, 413(6858), 844–848.

Watanabe, T. & Sasaki, Y. (2015). Perceptual learning: Toward a comprehensive theory. *Annual review of psychology*, 3(66), 197–221.

Watson, A. B. & Fitzhugh, A. (1990). The method of constant stimuli is inefficient. *Perception & Psychophysics*, 47(1), 87–91.

Watson, A. B. & Pelli, D. G. (1983). QUEST: A Bayesian adaptive psychometric method. *Attention, Perception, & Psychophysics*, 33(2), 113–120.

Webster, K. E., Edwin Dickinson, J., Battista, J., McKendrick, A. M., & Badcock, D. R. (2011). Increased internal noise cannot account for motion coherence processing deficits in migraine. *Cephalalgia*, 31(11), 1199–1210.

Wei, X. X. & Stocker, A. A. (2012). Efficient coding provides a direct link between prior and likelihood in perceptual Bayesian inference. *Advances in Neural Information Processing Systems*, 25(May), 1–9.

Weiss, Y., Edelman, S., & Fahle, M. (1993). Models of Perceptual-Learning in Vernier Hyperacuity. *Neural Comput.*, 5(5), 695–718.

Westheimer, G., Gilbert, C. D., Kapadia, M., & Ito, M. (2000). Interactions between attention, context and learning in primary visual cortex. *Vision Res*, 40(10-12), 1217–26.

Wichmann, F. A. & Hill, N. (2001a). The psychometric function: II. Boostrap-based confidence intervals and sampling. *Perception & Psychophysics*, 63(8), 1314–1329.

Wichmann, F. A. & Hill, N. J. (2001b). The psychometric function:{I}. {Fitting}, sampling, and goodness of fit. *Perception & psychophysics*, 63(8), 1239–1313.

Wiesel, T. N. & Hubel, D. H. (1963). Responses in Striate Deprived of Vision Cortex of One Eye. *Journal of Neurophysiology*, 26(6), 1003–1017.

Wiesel, T. N. & Hubel, D. H. (1965). Extent of Recovery From the Effects of Visual Deprivation in Kittens. *Journal of Neurophysiology*, 28(6), 1060–1072.

Wiesel, T. N. & Hubel, D. H. (1966). Spatial and chromatic interactions in the lateral geniculate body of the rhesus monkey. *Journal of Neurophysiology*, 29(6), 1115–1156.

Wiesel, T. N., Hubel, D. H., & others (1963). Effects of visual deprivation on morphology and physiology of cells in the cat's lateral geniculate body. *J Neurophysiol*, 26(978), 6.

Wilcox, L. M. & Allison, R. S. (2009). Coarse-fine dichotomies in human stereopsis. *Vision Research*, 49(22), 2653–2665.

Wilcox, L. M. & Hess, R. F. (1995). D max for stereopsis depends on size, not spatial frequency content. *Vision research*, 35(8), 1061–1069.

Wilcox, L. M. & Hess, R. F. (1996). Is the site of non-linear filtering in stereopsis before or after binocular combination? *Vision Research*, 36(3), 391–399.

Wilcox, L. M. & Hess, R. F. (1997). Scale selection for second-order (non-linear) stereopsis. *Vision Research*, 37(21), 2981–2992.

Wilkinson, F. & Crotogino, J. (2000). Orientation discrimination thresholds in migraine: a measure of visual cortical inhibition. *Cephalalgia*, 20(1), 57–66.

Williams, D. W. & Sekuler, R. (1984). Coherent global motion percepts from stochastic local motions. *Vision Research*, 24(1), 55–62.

Wilmer, J. B. (2008). How to use individual differences to isolate functional organization, biology, and utility of visual functions; With illustrative proposals for stereopsis. *Spatial Vision*, 21(6), 561–579.

Wilson, D. A. & Stevenson, R. J. (2003). Olfactory perceptual learning: The critical role of memory in odor discrimination. *Neuroscience and Biobehavioral Reviews*, 27(4), 307–328.

Wilson, H. R., Ferrera, V. P., & Yo, C. (1992). A Psychophysically Motivated Model for 2-Dimensional Motion Perception. *Visual Neuroscience*, 9(1), 79–97.

Wilson, H. R. & Wilkinson, F. (1998). Detection of global structure in Glass patterns: Implications for form vision. *Vision Research*, 38(19), 2933–2947.

Wilson, W. J. (2012). The Rescorla-Wagner Model , Simplified.

Xiao, L. Q., Zhang, J. Y., Wang, R., Klein, S. A., Levi, D. M., & Yu, C. (2008). Complete Transfer of Perceptual Learning across Retinal Locations Enabled by Double Training. *Current Biology*, 18(24), 1922–1926.

Xu, X., Ichida, J. M., Allison, J. D., Boyd, J. D., Bonds, a. B., & Casagrande, V. a. (2001). A comparison of koniocellular, magnocellular and parvocellular receptive field properties in the lateral geniculate nucleus of the owl monkey (Aotus trivirgatus) Xiangmin Xu, Jennifer M. Ichida, John D. Allison, Jamie D. Boyd, A. B. Bonds and Vivien A. *Journal of Physiology*, 531(1), 203–218.

Yang, T. & Maunsell, J. H. R. (2004). The Effect of Perceptual Learning on Neuronal Responses in Monkey Visual Area V4. *Journal of Neuroscience*, 24(7), 1617–1626.

Yang, T. & Shadlen, M. N. (2007). Probabilistic reasoning by neurons. *Nature*, 447(7148), 1075–1080.

Yu, C., Klein, S. A., & Levi, D. M. (2004). Perceptual learning in contrast discrimination and the (minimal) role of context. *Journal of Vision*, 4(2004), 169–182.

Zednik, C. (2018). Mechanisms in Cognitive Science. In S. Glennan & P. Illaria (Eds.), *The Routledge Handbook of Mechanisms and Mechanical Philosophy* chapter 29, (pp. 389–400). Oxon: Routledge, 1st edition.

Zeki, S. (1974). Functional organization of a visual area in the posterior bank of the superior temporal sulcus of the rhesus monkey. *The Journal of Physiology*, 236(3), 549–573.

Zeki, S. (1978). Functional specialisation in the visual cortex of the rhesus monkey. *Nature*, 274(5670), 423–428.

Zeki, S. (1983). Colour coding in the cerebral cortex: The reaction of cells in monkey visual cortex to wavelengths and colours. *Neuroscience*, 9(4), 741–765.

Zeki, S. (2003). Functional Specialisation in the Visual Brain: Probable and Improbable Visual Areas. In R. S. Frackowiak, K. J. Friston, C. D. Frith, R. J. Dolan, C. J. Price, S. Zeki, J. T. Ashburner, & W. D. Penny (Eds.), *Human Brain Function: Second Edition* chapter 9, (pp. 161–170). Academic Press, 2nd edition.

Zhang, J., Zhang, G., Xiao, L., Klein, S. A., Levi, D. M., & Yu, C. (2010). Rule-Based Learning Explains Visual Perceptual Learning and Its Specificity and Transfer. *Journal of Neuroscience*, 30(37), 12323–12328.

Zhang, J. Y., Kuai, S. G., Xiao, L. Q., Klein, S. A., Levi, D. M., & Yu, C. (2008). Stimulus coding rules for perceptual learning. *PLoS Biology*, 6(8), 1651–1660.

Zhaoping, L. (2002). Pre–attentive segmentation and correspondence in stereo. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 357(1428), 1877–1883.

Zhaoping, L. (2005). Border ownership from intracortical interactions in visual area V2. *Neuron*, 47(1), 143–153.

Zhaoping, L. (2014). *Understanding vision: Theory, models, and data.* Oxford University Press, USA.

Zhaoping, L. (2017). Feedback from higher to lower visual areas for visual recognition may be weaker in the periphery: Glimpses from the perception of brief dichoptic stimuli. *Vision Research*, 136, 32–49.

Zhaoping, L. & Ackermann, J. (2017). Reversed depth in anti-correlated random dot stereograms and central-peripheral dif- ference in visual inference. *bioRxiv preprint*, (pp. 1–8).

Zhou, J., Liu, R., Zhou, Y., & Hess, R. F. (2014). Binocular combination of second-order stimuli. *PLoS ONE*, 9(1), e84632.

Ziegler, L. R. & Hess, R. F. (1999). Stereoscopic depth but not shape perception from second-order stimuli. *Vision Research*, 39(8), 1491–1507.

Zohary, E., Celebrini, S., Britten, K. H., & Newsome, W. T. (1994). Neuronal Plasticity that Underlies Improvement in Perceptual Performance. *Science*, 263(5151), 1289–1292.

Zylberberg, A., Roelfsema, P. R., & Sigman, M. (2014). Variance misperception explains illusions of confidence in simple perceptual decisions. *Consciousness and Cognition*, 27(1), 246–253.